HAL open science

Les syntagmes nominaux polydéfinis en romani d'Albanie.

Aurore Tirard

▶ To cite this version:

Aurore Tirard. Les syntagmes nominaux polydéfinis en romani d'Albanie.. Linguistique. Université Clermont Auvergne [2017-2020], 2019. Français. NNT : 2019CLFAL019 . tel-03336228

HAL Id: tel-03336228

https://theses.hal.science/tel-03336228v1

Submitted on 7 Sep 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Université Clermont Auvergne
École doctorale n°370 *Lettres, Sciences Humaines et Sociales*
Unité de Recherche EA n°999 *Laboratoire de Recherche sur le Langage*

Les syntagmes nominaux polydéfinis en romani d'Albanie

Aurore TIRARD

Thèse de doctorat en Sciences du langage

dirigée par **Friederike SPITZL-DUPIC** & **Hana GRUET-ŠKRABALOVA**

Soutenue le 12 décembre 2019

Jury

Gerda **HAßLER**	Professeure, Universität Potsdam – présidente du jury, pré-rapporteuse
Victor **FRIEDMAN**	Professeur émérite, The University of Chicago – pré-rapporteur
Benjamin **FAGARD**	Chargé de recherche habilité, C.N.R.S. LATTICE & E.N.S. – pré-rapporteur
Patricia **CABREDO HOFHERR**	Chargée de recherche habilitée, C.N.R.S. S.F.L. & Université Paris 8
Friederike **SPITZL-DUPIC**	Professeure, Université Clermont Auvergne – directrice de recherche
Hana **GRUET-ŠKRABALOVA**	Maîtresse de conférences, Université Clermont Auvergne – encadrante

Résumé

Cette thèse porte sur les syntagmes nominaux comportant un adjectif épithète articulé en romani d'Albanie. Ce phénomène consiste en l'ajout d'un article défini « adjectival » au sein du syntagme nominal et rappelle le *polydefinite NP* ou *determiner spreading* grec. Nous montrons qu'il correspond à la détermination multiple de type *pattern* syntaxique, selon la typologie d'Alexiadou (2014), et au type 1 des langues selon celle de Lekakou et Szendrői (2012). Cette structure remonte probablement au vieux romani, ce qui implique que toutes les variétés actuelles l'aient un jour connue. Elle provient du contact avec le grec par convergence de type *pattern replication* (Matras 2009). Pour étudier les syntagmes polydéfinis, nous avons réalisé deux missions de terrain à Korçë en 2013 et 2014, afin d'y recueillir un corpus à la fois spontané et semi-spontané, et de combiner analyse qualitative et quantitative. Ils s'emploient surtout avec des adjectifs prédicatifs, restrictifs et intersectifs, informationnellement nouveaux et placés après le nom. Le déterminant « nominal » est l'article défini, le possessif ou le démonstratif ; le déterminant « adjectival » est l'article défini qui joue un rôle d'opérateur de focalisation. Nos résultats montrent que les syntagmes polydéfinis font l'objet de « contraintes préférentielles » (Thuilier 2012). La partie adjectivale des syntagmes polydéfinis est un cas de « modification indirecte » (Cinque 2010) et l'emplacement post-nominal est une « post-zone » (Nølke 2001). Parce qu'il est le lieu du focus interne, il a plusieurs fonctions : identification restrictive du référent, contraste, emphase, récapitulation, *topic-shift*, création de surnoms et « romanisation » des expressions empruntées. Les syntagmes polydéfinis et la place de l'adjectif sont sujets à variation dialectale, diachronique, diagénique et diaphasique en romani d'Albanie. Nous avons constaté qu'un changement de l'ordre des constituants nominaux s'est produit ou est en cours. L'ordre des mots hérité est DEF A N tandis que celui innovant est DEF N A. Les syntagmes polydéfinis DEF N DEF A ont servi d'intermédiaire d'un ordre à l'autre, puisqu'il s'agit d'une structure ancienne mais à adjectif postposé. À l'échelle du romani albanais, on peut dire que la postposition de l'adjectif est devenue majoritaire en tous syntagmes nominaux.

Mots-clé : variation linguistique, changement linguistique, syntaxe, syntagme nominal, ordre des mots, définitude, surdétermination, focalisation, contact de langues, linguistique de terrain, romani, grec, diachronie, aire linguistique, Albanie, Balkans.

Abstract

This dissertation is devoted to the study of nominal phrases with an articulated attributive adjective in Albanian Romani. This phenomenon consists in the use of an “adjectival” definite article within the nominal phrase and reminds of the *polydefinite NPs* or the *Determiner Spreading* in Greek. It corresponds to the *syntactic pattern* kind of multiple determination according to Alexiadou (2014) typology, and to the type 1 langages according to Lekakou and Szendrői (2012) typology. It goes probably back to Early Romani, which means that all current Romani varieties once had it. It arises from contact with Greek through *pattern replication* (Matras 2009). I have done two fieldworks in Korça in 2013 and 2014, where I collected a corpus of spontaneous and semi-spontaneous speech to combine qualitative and quantitative analysis. The structure is mainly used with predicative, restrictive and intersective adjectives, new in the information structure and postponed to the noun. The “nominal” determiner is the definite article, the possessive or demonstrative; the “adjectival” determiner is the definite article with the function of focalisation operator. The results show that polydefinite phrases are subject to “preferential constraints” (Thuilier 2012). The adjectival subpart of the phrase is an “indirect modification” (Cinque 2010) and the postnominal slot is a “post-zone” (Nølke 2001). Since it is the locus of internal focus, it has several functions: restrictive identification of the referent, contrast, emphasis, synthesis, topic-shift, nickname creation and “romanisation” of loanwords. The polydefinite phrases and the adjectival place undergo dialectal, diachronic, diagenic and diaphasic variation in Albanian Romani. I have observed that a change in the nominal constituents order has taken place or is ongoing. The inherited word order is D A N while the new word order ist D N A. The word order D N D A is an intermediate solution and was used as a bridge from one to another as it is an ancient structure but with postposed adjective. With respect to Albanian Romani we demonstrate that adjective postposition has become predominant for all nominal phrases.

Keywords: linguistic variation, linguistic change, syntax, nominal phrase, word order, definiteness, overdetermination, focus, language contact, linguistic fieldwork, Romani, Greek, diachrony, linguistic area, Albania, Balkans.

Remerciements

Mes deux missions de terrain en 2013 et 2014 ont été rendues possibles grâce au soutien financier de l'École Doctorale n°265 *Langues, littératures et sociétés du monde* et à celui de l'Unité Mixte de Recherche n°8202 SeDyL (*Structure et Dynamique des Langues*). L'École Normale Supérieure de Paris a permis le financement des premières années de ce travail grâce au contrat doctoral effectué à l'Inalco (*Institut National des Langues et Civilisations Orientales*). C'est l'EA 99 LRL (*Laboratoire de Recherche sur le Langage*) qui a permis ma soutenance.

La thèse n'aurait pu s'achever sans l'aide de mes deux directrices de recherche, Friederike Spitzl-Dupic et Hana Gruet-Škrabalová (Université Blaise Pascal / Clermont Auvergne), qui ont montré leur soutien indéfectible dans le parcours semé d'embûches du doctorat, notamment lors de la phase finale de rédaction. J'adresse également mes remerciements à mes précédents directeurs, Marcel Courthiade et Sophie Vassilaki (Inalco), grâce à qui j'ai pu affiner mon travail de recherche sur l'aire linguistique balkanique. Je suis également très reconnaissante à mon ancienne directrice Evangelia Adamou (UMR 7107 Lacito, *Langues et Civilisations à Tradition Orale*, du CNRS) de m'avoir suggéré d'effectuer un test linguistique puis aidée à le concevoir, m'amenant ainsi à la collecte du corpus semi-spontané et à son étude quantitative.

De plus, je remercie le jury qui a évalué mon travail : Gerda Haßler (Universität Potsdam), Victor A. Friedman (The University of Chicago), Benjamin Fagard (UMR 8094 Lattice, *Langues, Textes, Traitements informatiques et Cognition*, du CNRS), Patricia Cabredo Hofherr (UMR 7023 SFL, *Structures Formelles du Langage*, du CNRS), Friederike Spitzl-Dupic et Hana Gruet-Škrabalová. Je n'oublie pas mon jury de soutenance à mi-parcours, qui a par deux fois évalué mon travail et m'a apporté de précieux conseils : Martine Vanhove (UMR 8135 LLACAN, *Langage, Langues et Cultures d'Afrique*), Patricia Cabredo Hofherr, Sophie Vassilaki et Evangelia Adamou. Je tiens à remercier Elisabeth Harriet et Paul Lotin (LRL) qui ont contibué à ma soutenance sur les plans pratique et technique.

En ce qui concerne mes missions de terrain, je tiens à témoigner ma profonde gratitude à mes informateurs et collaborateurs roms, qui ont su m'apporter le soutien nécessaire, tant sur le plan pratique que relationnel, pour le bon déroulement de mes séjours et de mes recherches. Je pense particulièrement à mon hôte de Korçë, Arben Kosturi, avec qui mon enseignant de langue et civilisation rromani, Marcel Courthiade, m'a mise en contact et auprès duquel il m'a recommandée. L'accueil qu'Arben et sa famille m'ont réservé a été chaleureux, bienveillant et généreux : ils m'ont aussitôt acceptée et intégrée à leur famille. Je souhaite, par le présent travail, les en remercier tous chaleureusement. Je remercie profondément Orhan Galjuš et Fenia Papakonstantinou, Evangelia Adamou et Dimitris Katsavos, qui m'ont permis de poursuivre mes recherches en Grèce malgré les obstacles auxquels j'ai été confrontée en 2014.

Pour avoir contribué à alimenter mes réflexions sur mon sujet d'étude, je tiens à remercier les nombreux collègues et enseignants avec qui j'ai eu des discussions très fructueuses sur mes recherches. Il s'agit entre autres de Yaron Matras et Viktor D. Leggio (The University of Manchester),

Isabelle Léglise, Claudine Chamoreau et Anaïd Donabedian (SeDyL), François Jacquesson, Zlatka Gentcheva, Isabelle Bril, Lameen Souag, Catherine Taine-Cheikh et Pablo Kirtchuk-Halevi (Lacito), Alexandre François (Lacito/Lattice), Lidia Lebas-Fraczak, Hana Gruet-Škrabalová, Ciara Wigham et Véronique Quanquin (Université Clermont Auvergne), Walter Breu (Universität Konstanz) et Isabelle Buchstaller (Universität Leipzig / Universität Duisburg Essen).

J'adresse mes plus chaleureux remerciements à mes fidèles relecteurs et correcteurs Elisabeth Laroche, Michel Tirard et Marco Zito. Un grand merci aux collègues et amis qui ont accepté de relire certaines sections de mon travail et de m'adresser leurs conseils avisés : Victor A. Friedman, Bernard Lory, Marie-Laure Coppolani, Rola Skaf, Christine Salastri et Apolline Lehmann.

Je tiens à saluer les collègues aux côtés desquels j'ai effectué mes recherches et enseigné durant mon contrat doctoral et mes contrats successifs d'ATER. J'ai eu la chance de mener avec eux de passionnantes discussions non seulement sur mon sujet d'étude, mais aussi sur bien d'autres phénomènes linguistiques intrigants. Il s'agit entre autres de Snežana Gadževa, Alexandru Mardale, Anton Antonov, Sophie Vassilaki, Christine Bonnot, Annie Montaut, Henri Menantaud (†) et Thomas Szende (Inalco), Lidia Lebas-Fraczak et Hana Gruet-Škrabalová (Université Clermont Auvergne), Michel Lecolle, Anne Didelot et Annabelle Seoane (Université de Lorraine), Marie-Armelle Camussi-Ni, Malo Morvan et Diana Andrei (Université Rennes 2). Je suis également reconnaissante aux collègues ingénieurs et techniciens Séverine Guillaume, Anne Behaghel-Dindorf, Jérôme Picard (Lacito), Anne Bélingard (Lacito/Centre de documentation A.G. Haudricourt), Fabien Dugal (UPS 2259), Guy Sinelle (Inalco).

J'ai une pensée toute particulière pour les doctorants qui ont fait une partie de leur parcours de thèse en même temps que moi : Marie-Laure Coppolani, Tom Durand, Santiago Sánchez Moreano, Jhonnatan Rangel Murueta, Jean-François Nunez et Jeanne Gamonet (SeDyL), Rola Skaf (Llacan), Grégory Nardozza (†) (Paris IV), Emilie Riguel (Paris III), Camille Simon, Karell Marchand, Benjamin Touati, Aïcha Saad, Komi Simnara, Evgeniya Gutova, Cristian Pădure et Asuka Matsumoto (Lacito), Siglinde Pape, Frédéric Mourier, Aline Auriel, Laurent Mario et Zulimar Colina (LRL, UCA), Gisèle Casagranda, Feliciano José Pedro et Hadria Harket-Khelifi, (EA 3476 CREM, *Centre de Recherche sur les Médiations*, Université de Lorraine).

Enfin, je tiens à saluer mes hôtes albanais Arben Kosturi, Moza Kosturi, Marinela Kosturi, Enxhellushë Kosturi, Brikena Kosturi et Denis Kosturi, mes hôtes serbes Ismet Jašarević, Kimet Jašarević, Amir Jašarević et Ramajana Jašarević, mes informateurs de Grèce Nikos Paiteris et Toni. Je salue également mes amis roms Jeanne Gamonet, Saimir Mile, Orhan Galjuš, Cristian Pădure, Ionela Pădure, Slavka Stefanova, Mitko Angelov, Fatos Koçi, Emiliano Aliu, Françesko Manaj, Danjel Hyseni, Manjola Veizi, Marjus Veizi, Gurali Mejdani, Besmir Dalipi, Bledi Taho, ainsi que Walter de Maggio, Artan Grabocka, et tous les autres.

Table des matières

Résumé I

Abstract II

Remerciements III

Table des matières V

Table des figures XV

Table des tableaux XVIII

Table des abréviations XXII

Introduction 1

Chapitre I : Les Roms d'Albanie et leur langue 5

1. Les Roms d'Albanie 5

1.1. Éléments démographiques 5

1.1.1. Taille de la population 5

1.1.2. Répartition géographique de la population 7

1.2. Éléments sociologiques 9

1.2.1. Instruction 9

1.2.2. Activité professionnelle 11

1.2.3. Discrimination 12

1.2.4. Korçë 14

1.3. Les groupes 16

1.3.1. Ethnonymes 17

1.3.2. Mečkars 21

1.3.3. Arlis 21

1.3.4. Čergars I et II 23

1.3.5. Bamils 24

1.3.6. Kurtofis 25

1.3.7. Différences entre les groupes 25

1.4. Roms et Égyptiens 26

2. La langue romani 29

2.1. Dialectologie 29

2.1.1. Spécificités de la langue 29

2.1.2. Modèles dialectologiques 30

2.1.3. Classification dialectologique 32

2.2. Le contact de langues 34

2.2.1. En diachronie 34

2.2.2. En synchronie 35

2.3. Les variétés des Balkans 37

2.4. Les variétés d'Albanie 40

2.4.1. Mečkar 40

2.4.2. Arli 41

2.4.3. Čergar I (vlax) 41

2.4.4. Čergar II (non-vlax) *42*
2.4.5. Bamil et kurtofi *42*
2.4.6. Les néo-locuteurs *42*
2.4.7. Contact de langues *43*

3. Méthodologie générale 44

3.1. Questionnaire et élicitation ? 44
3.2. Linguistique de corpus 48
3.3. Linguistique de terrain 50
3.3.1. La question de la maîtrise de la langue *50*
3.3.2. Le déroulement des séjours *50*
3.3.3. Le rôle de chercheuse sur le terrain *52*
3.4. Linguistique de données « naturalistes » 53
3.5. Linguistique de données expérimentales 55
3.6. Triangulation 56

..... Chapitre II : Théorie 59

1. La définitude 59

1.1. Définition 59
1.1.1. Présupposition d'unicité et d'existence *59*
1.1.2. Expressions définies complètes et incomplètes *61*
1.1.3. La catégorie grammaticale *62*
1.2. Effet sur la structure de l'information 64
1.3. L'article défini 64
1.3.1. Définition *64*
1.3.2. Typologie *65*
1.3.3. Fonctions en diachronie *67*
1.3.4. Fonctions en synchronie *68*
1.3.5. Usage avec les noms propres *70*

2. L'adjectif 72

2.1. Définition 72
2.2. Aspects syntaxiques 73
2.2.1. Fonctions *73*
2.2.2. Distribution *74*
2.2.3. Ordre des adjectifs *77*
2.3. Aspects sémantiques 82
2.3.1. Classes d'adjectifs *82*
2.3.2. Interprétation du syntagme nominal *84*

3. La structure de l'information 89

3.1. Structure et statut de l'information 89
3.2. Focus et topique 91
3.3. La focalisation 93
3.4. Deux types de focus 95
3.5. Dans le syntagme nominal 97

..... Chapitre III : État de l'art 103

1. Le syntagme nominal en romani 103

1.1. L'ordre des constituants 103
1.2. Le nom 104
1.2.1. Catégories 104
1.2.2. Flexion 106
1.3. Le déterminant 107
1.3.1. Article défini 107
1.3.1.1. Origine 107
1.3.1.2. Flexion 109
1.3.1.3. Fusion avec certaines prépositions 109
1.3.1.4. Distribution 111
1.3.1.5. Fonctions 111
1.3.1.6. Usage avec les noms propres 116
1.3.2. Article indéfini 116
1.3.2.1. Jekh 117
1.3.2.2. Article zéro 119
1.3.3. Démonstratif 122
1.3.3.1. Formes et emplois 122
1.3.3.2. Distribution et double détermination 122
1.3.4. Possessif 125
1.3.4.1. Formes 125
1.3.4.2. Distribution et double détermination 126
1.3.5. Quantifieur et numéral cardinal 127
1.3.5.1. Formes 127
1.3.5.2. Distribution et double détermination 128
1.3.6. Syntagme génitival 130
1.3.6.1. Forme 130
1.3.6.2. Fonction de modifieur adnominal 131
1.3.6.3. Fonction de déterminant 134
1.4. L'adjectif 136
1.4.1. Flexion 136
1.4.2. Degrés 137
1.4.2.1. Comparatif 137
1.4.2.2. Superlatif 138
1.4.3. Distribution 138
1.4.3.1. Antéposition canonique 138
1.4.3.2. Pourquoi postposer l'adjectif ? 139

2. La répétition de l'article défini 142

2.1. En romani 142
2.2. En grec 147
2.2.1. Similarité du grec et du romani 147
2.2.2. Les constituants du syntagme polydéfini en grec 149
2.2.2.1. Nom 149
2.2.2.2. Adjectif 149
2.2.2.3. Déterminant 150
2.2.2.4. Ordre des constituants 158
2.2.2.5. Ordre des adjectifs 160
2.3. Analyse syntaxique 163
2.3.1. Accord en définitude 163
2.3.2. Proposition relative sous-jacente 164
2.3.3. Ellipse du nom 166
2.3.4. Apposition 169

2.3.4.1. Deux types d'apposition 169
2.3.4.2. Apposition étroite 173
2.3.4.3. Apposition lâche 176
2.4. Analyse sémantique 178
2.4.1. … de l'article : la surdétermination 178
2.4.2. … de l'adjectif : l'intersectivité et la restrictivité 179
2.4.3. … de l'adjectif : les exclus de la polydéfinitude 185
2.4.4. … de l'adjectif : les exceptions 188
2.5. Analyse informationnelle 192
2.6. Bilan 196
2.6.1. La polydéfinitude en grec 196
2.6.2. Problématique 198

....Chapitre IV : Analyse du corpus spontané 201

1. Méthodologie 201

1.1. Élaboration du corpus spontané 201
1.2. Les informateurs 202
1.3. L'usage du romani au quotidien 203
1.3.1. Dans la famille 203
1.3.2. Hors de la famille 205
1.4. Le matériel et les données collectées 206
1.5. La transcription et la présentation des exemples 208
1.6. Analyse des données 209

2. Syntagmes sans rupture interne 211

2.1. Syntagmes polydéfinis 211
2.1.1. DEF N DEF A 211
2.1.1.1. Effet de sélection 211
2.1.1.2. Effet d'emphase 218
2.1.1.3. Effet hybride 219
2.1.2. DEF A DEF N 223
2.1.3. Synthèse 224
2.1.3.1. Bilan 224
2.1.3.2. Focus 225
2.1.3.3. Place de l'adjectif 227
2.2. Syntagmes monodéfinis 227
2.2.1. DEF A N 228
2.2.2. DEF N A 231
2.2.3. Synthèse 235
2.2.3.1. Bilan 235
2.2.3.2. Focus 235
2.2.3.3. Place de l'adjectif 236
2.3. Synthèse générale 239
2.3.1. Type d'adjectif 239
2.3.2. Modification directe et indirecte 240
2.3.3. Place de l'adjectif 242
2.3.4. Place de l'adjectif et variation dialectale 243
2.3.5. Pourquoi un deuxième article défini ? 250
2.3.6. Réponse à la problématique 255

3. Syntagmes avec rupture interne 263

3.1. Syntagmes polydéfinis 263
3.1.1. DEF N/ DEF A *263*
3.1.1.1. Cas sans insertion 264
3.1.1.2. Cas avec insertion 265
3.1.1.3. Les « faux amis » def n et dem n 267
3.1.2. DEF A/ DEF N *269*
3.1.3. Synthèse *269*
3.1.3.1. Bilan 269
3.1.3.2. Focus 270
3.2. Syntagmes monodéfinis 270
3.5.1. DEF A/ N *270*
3.2.1. DEF N/ A *271*
3.2.2. Synthèse *272*
3.2.2.1. Bilan 272
3.2.2.2. Focus 273
3.3. Synthèse générale 273
3.3.1. Type d'adjectif *274*
3.3.2. Modification directe et indirecte *274*
3.3.3. Place de l'adjectif et variation dialectale *275*
3.3.4. Rôle de la rupture et de l'insertion *276*
3.3.5. Réponse à la problématique *277*

Chapitre V : Analyse du corpus semi-spontané 287

1. Méthodologie 287

1.1 Les tests réalisés 287
1.1.1 Introduction *287*
1.1.2 Test 1 ..... *289*
1.1.2.1 Principe 289
1.1.2.2 Matériel 290
1.1.2.3 Procédure 291
1.1.2.4 Enregistrement 292
1.1.2.5 Participants 292
1.1.2.6 Difficultés rencontrées 293
1.1.3 Test 2 ..... *294*
1.1.3.1 Procédure 294
1.1.3.2 Matériel 295
1.1.3.3 Enregistrement 296
1.1.3.4 Participants 296
1.1.3.5 Difficultés rencontrées 297
1.2 Dépouillement et codage 298

2. Résultats : variables linguistiques 299

2.1 Distribution des syntagmes définis 299
2.1.1 Test 1 ..... *300*
2.1.2 Test 2 ..... *300*
2.1.3 Tests 1 et 2 ..... *301*
2.1.4 Témoins *302*
2.1.5 Synthèse *302*
2.2 Variable 1 : DEFINITUDE 302
2.2.1 Quel déterminant répéter ? *303*
2.2.1.1 Tests 1 et 2 303

2.2.1.2 Témoins 303
2.2.2 Répéter l'article défini en présence d'un autre déterminant 304
2.2.2.1 Tests 1 et 2 304
2.2.2.2 Témoins 305
2.2.2.3 Synthèse 305
2.3 Variable 2 : ORDRE DES CONSTITUANTS 306
2.3.1 Syntagmes monodéfinis 306
2.3.1.1 Tests 1 et 2 306
2.3.1.2 Témoins 309
2.3.1.3 Répétition de l'affixe de cas 310
2.3.2 Syntagmes polydéfinis 311
2.3.2.1 Un adjectif 311
2.3.2.2 Plusieurs adjectifs 312
2.3.2.3 Répétition de l'affixe casuel 319
2.3.2.4 Répétition de la préposition 320
2.3.2.5 Témoins 324
2.3.3 Syntagmes introduits par un autre déterminant 325
2.3.3.1 Tests 1 et 2 325
2.3.3.2 Témoins 326
2.3.4 Synthèse 326
2.4 Variable 3 : TYPE D'ADJECTIF 329
2.4.1 Poids des adjectifs et type syntaxique 331
2.4.2 Origine et sémantique des adjectifs dans les syntagmes polydéfinis 332
2.4.3 Origine des adjectifs dans tous les syntagmes définis 335
2.4.4 Sémantique des adjectifs dans tous les syntagmes définis 342
2.4.5 La combinaison sémantique adjectif – nom 345
2.4.6 Témoins 347
2.4.7 Synthèse 350
2.5 Synthèse des variables linguistiques 355

3. Résultats : variables sociales 358

3.1 Introduction 358
3.1.1 Variantes étudiées 358
3.1.2 Variables étudiées 358
3.1.3 Le genre 361
3.1.4 Le système scolaire en Albanie 361
3.1.5 Le militantisme 362
3.1.6 Les variétés 363
3.2 Variable 1 : GENRE 364
3.2.1 Participants albanais 364
3.2.1.1 Structure des syntagmes nominaux définis 364
3.2.1.2 Proportion d'adjectifs anté- et postposés 365
3.2.2 Témoins 366
3.2.2.1 Structure des syntagmes nominaux définis 366
3.2.2.2 Proportion d'adjectifs anté- et postposés 367
3.3 Variable 2 : CLASSE D'AGE 368
3.3.1 Participants albanais 368
3.3.1.1 Structure des syntagmes nominaux définis 368
3.3.1.2 Proportion d'adjectifs anté- et postposés 369
3.3.2 Témoins 369
3.3.2.1 Structure des syntagmes nominaux définis 369
3.3.2.2 Proportion d'adjectifs anté- et postposés 370

3.4 Variable 3 : EDUCATION 371
3.4.1 Participants albanais 371
3.4.1.1 Structure des syntagmes nominaux définis 371
3.4.1.2 Proportion d'adjectifs anté- et postposés 372
3.4.2 Témoins 372
3.4.2.1 Structure des syntagmes nominaux définis 372
3.4.2.2 Proportion d'adjectifs anté- et postposés 373
3.5 Variable 4 : MILITANTISME 374
3.5.1 Participants albanais 374
3.5.1.1 Structure des syntagmes nominaux définis 374
3.5.1.2 Proportion d'adjectifs anté- et postposés 375
3.5.2 Témoins 376
3.5.2.1 Structure des syntagmes nominaux définis 376
3.5.2.2 Proportion d'adjectifs anté- et postposés 377
3.6 Variable 5 : VARIETE 378
3.6.1 Participants albanais 378
3.6.1.1 Structure des syntagmes nominaux définis 378
3.6.1.2 Proportion d'adjectifs anté- et postposés 379
3.6.2 Témoins 380
3.6.2.1 Structure des syntagmes nominaux définis 380
3.6.2.2 Proportion d'adjectifs anté- et postposés 381
3.7 Synthèse des variables sociales 381
3.7.1 Test 1 ~ test 2 381
3.7.2 Hiérarchie des variables 382
3.7.3 Profil des locuteurs 383
3.7.3.1 Participants albanais 383
3.7.3.2 Témoins 384
3.7.3.3 Regroupements de profils 384
3.7.4 Perspectives 385
3.7.4.1 Difficultés 385
3.7.4.2 Solutions 386
3.7.4.3 Variable 6 : INDIVIDU 388
3.7.5 Synthèse 393

4. Synthèse générale 394

Chapitre VI : Discussion 407

1. Origine de la polydéfinitude 407

1.1 Problématique 407
1.2 Histoire : le romani albanais ancien 408
1.3 Géographie : le romani des autres régions 411
1.3.1. Descriptions d'autres variétés 411
1.3.2. Le corpus RMS 411
1.3.3. Données roumaines contemporaines 415
1.3.4. Données roumaines anciennes 416
1.3.5. Autres données 420
1.3.5.1 Macédoine et Kosovo 420
1.3.5.2 Grèce 420
1.3.5.3 Bulgarie 423
1.3.5.4 Hongrie 423
1.4 Scénario géographique 424
1.4.1. Dans les Balkans 424

1.4.1.1 Le romani 424
1.4.1.2 Les autres langues....425
1.4.1.2.1 Article défini 426
1.4.1.2.2 « Double » article défini 427
1.4.1.2.3 Quelle langue d'emprunt ?....434
1.4.2. Hors des Balkans 436
1.4.2.1 Le romani 436
1.4.2.2 Les autres langues....438
1.5 Scénario historique 444

2. Origine de la postposition de l'adjectif 446

2.1 Problématique 446
2.2 Histoire : le romani albanais ancien 447
2.3 Géographie : le romani des autres régions 447
2.3.1. Descriptions d'autres variétés 447
2.3.2. Données roumaines contemporaines....448
2.3.3. Données roumaines anciennes....448
2.3.4. Autres données....451
2.3.4.1. Macédoine et Kosovo 451
2.3.4.2. Grèce....451
2.4. Scénario géographique 452
2.4.1. Dans les Balkans....452
2.4.1.1. Le romani 452
2.4.1.2. Les autres langues....452
2.4.2. Hors des Balkans 455
2.4.2.1. Le romani 455
2.4.2.2. Les autres langues....455
2.5. Scénario historique 457
2.5.1. Problème de chronologie des acquisitions 458
2.5.2. Solution de chronologie....459
2.5.3. Objections....461

3. Scénario socio-linguistique 462

3.1. Types d'études 462
3.2. Types de résultats 462
3.3. Nos résultats 464
3.4. Interprétation de nos résultats 465
3.4.1. DEF A N 465
3.4.2. DEF N A 465
3.4.3. DEF N DEF A....467

4. Scénario théorique 469

4.1. Pourquoi est née cette structure ? 469
4.1.1. L'article défini supplémentaire....469
4.1.2. La place de l'adjectif....471
4.1.3. Les fonctions....475
4.1.4. L'optionnalité 477
4.2. Le processus de transfert linguistique 479
4.2.1. Le mécanisme de changement syntaxique....479
4.2.2. Le mécanisme à l'œuvre en romani....482
4.3. Synthèse....483

5. Réponse à la problématique 485

.....Conclusion 505

.....Bibliographie 513

.....Annexes.....539

1. Annexes du chapitre I : Les Roms d'Albanie et leur langue 539

2. Annexes du chapitre IV : Analyse du corpus spontané 542

2.1. Profil des locuteurs ayant prononcé les exemples spontanés 542
2.2. De la polysémie de certains termes de famille 544
2.3. Analyses 544
2.4. Corpus spontané : DEF N DEF A 549
2.5. Corpus spontané : DEF A DEF N 571
2.6. Corpus spontané : DEF A N 575
2.7. Corpus spontané : DEF N A 586
2.8. Corpus spontané défini et indéfini : autres exemples 594
2.9. Corpus spontané : DEF N/ DEF A.....599
2.10. Corpus spontané : DEF A ET DEM N 603
2.11. Corpus spontané : DEF N/ A 606

3. Annexes du chapitre V : Analyse du corpus semi-spontané 610

3.1 Métadonnées et codage 610
3.2 Variation linguistique – données 622
3.3 Variation sociale – tableaux 628

4. Annexes du chapitre VI : Discussion.....636

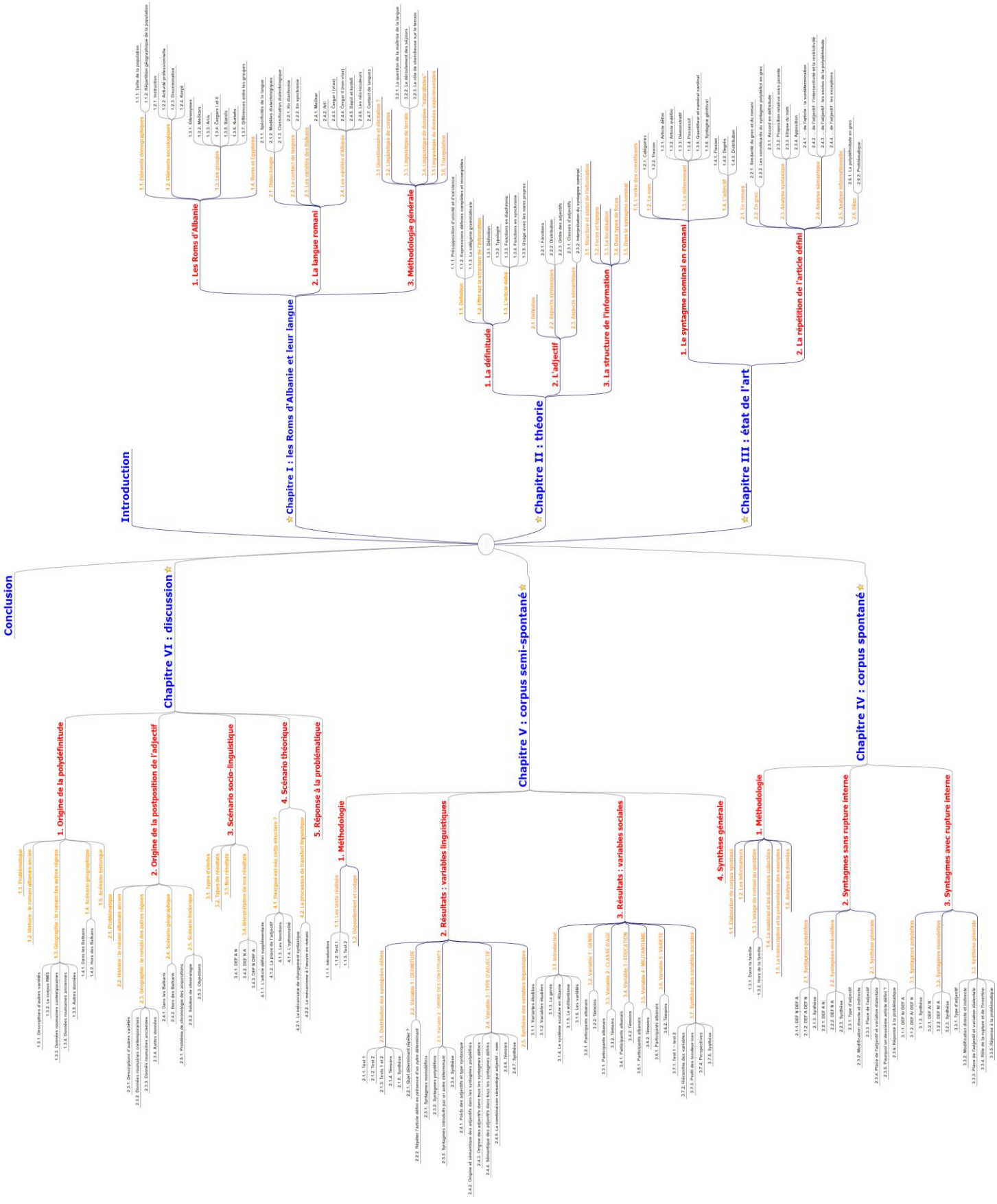

Introduction
Chapitre I : les Roms d'Albanie et leur langue
1. Les Roms d'Albanie
1.1. Démo-démographiques
1.1.1. Taille de la population
1.1.2. Répartition géographique de la population
1.2. Démo-socio-géographiques
1.2.1. Instruction
1.2.2. Activité professionnelle
1.2.3. Discrimination
1.2.4. Accès
1.3. Les groupes
1.3.1. Ethnonymes
1.3.2. Meshkors
1.3.3. Arlis
1.3.4. Cergars I et II
1.3.5. Gabels
1.3.6. Kortels
1.3.7. Différences entre les groupes
1.4. Roms et Egyptiens
2. La langue romani
2.1. Spécificités de la langue
2.1.1. Spécificités de la langue
2.1.2. Modèles dialectologiques
2.1.3. Classification dialectologique
2.2. Dialectologie
2.2.1. En diachronie
2.2.2. En synchronie
2.3. Le contact des langues
2.3.1. Le contact des langues
2.3.2. Le contact des langues
2.3.3. Les variétés des Balkans
2.4. Les variétés d'Albanie
2.4.1. Meshkor
2.4.2. Arli
2.4.3. Cergar I (Gjas)
2.4.4. Cergar II (non-Gjas)
2.4.5. Gabel de Kavajë
2.4.6. Kortel de Vlorë
2.4.7. Contact des langues
3. Méthodologie générale
3.1. Questionnaire de corpus
3.1.1. Questionnaire de corpus
3.1.2. Questionnaire de terrain
3.2. Le questionnaire de terrain
3.2.1. La question de la maîtrise de la langue
3.2.2. Le déroulement des séjours
3.2.3. Le rôle de chercheuse sur le terrain
3.3. Linguistique de données "naturalistes"
3.3.1. Linguistique des données expérimentales
3.3.2. Transposition
Chapitre II : théorie
1. La définitude
1.1. Définition
1.1.1. Présupposition d'unicité et d'existence
1.1.2. Expressions définies complètes et incomplètes
1.1.3. La catégorie grammaticale
1.2. Effet sur la structure de l'information
1.3. L'article défini
1.3.1. Définition
1.3.2. Typologie
1.3.3. Fonctions en diachronie
1.3.4. Fonctions en synchronie
1.3.5. Usage avec les noms propres
2. L'adjectif
2.1. Définition
2.2. Aspects syntaxiques
2.2.1. Fonctions
2.2.2. Distribution
2.2.3. Ordre des adjectifs
2.3. Aspects sémantiques
2.3.1. Classes d'adjectifs
2.3.2. Interprétation du syntagme nominal
3. La structure de l'information
3.1. Structure et statut de l'information
3.2. Focus et topique
3.3. La localisation
3.4. Deux types de focus
3.5. Dans le syntagme nominal
Chapitre III : état de l'art
1. Le syntagme nominal en romani
1.1. L'ordre des constituants
1.2. Le nom
1.2.1. Catégories
1.2.2. Thèmes
1.3. Le démonstratif
1.3.1. Article défini
1.3.2. Article indéfini
1.3.3. Démonstratif
1.3.4. Possessif
1.3.5. Quantifieur et numéral cardinal
1.3.6. Syntagme génitif
1.4. L'adjectif
1.4.1. Thèmes
1.4.2. Degrés
1.4.3. Distribution
2. La répétition de l'article défini
2.1. En romani
2.2. En grec
2.2.1. Similarité de grec et du romani
2.2.2. Les constituants du syntagme polydéfini en grec
2.3. Analyses syntaxiques
2.3.1. Accord en définitude
2.3.2. Préposition relative sous jonctif
2.3.3. Ellipse du nom
2.3.4. Apposition
2.4. Analyses sémantiques
2.4.1. ... de l'article : la sacralisation
2.4.2. ... de l'adjectif : l'intersectivité et la restrictivité
2.4.3. ... de l'adjectif : les excès de la polydéfinitude
2.4.4. ... de l'adjectif : les exceptions
2.5. Analyses informationnelles
2.5.1. La polydéfinitude en grec
2.5.2. Problématique
2.6. Bilan
Conclusion
Chapitre VI : discussion
1. Origine de la polyéfinitude
1.1. Problématique
1.2. Histoire : le romani albanais ancien
1.3. Géographie : le romani des autres régions
1.3.1. Descriptions d'autres variétés
1.3.2. Le corpus DMS
1.3.3. Données roumaines contemporaines
1.3.4. Données roumaines anciennes
1.3.5. Autres données
1.4. Dans les Balkans
1.4.1. Dans les Balkans
1.4.2. Hors des Balkans
1.5. Scénario historique
2. Origine de la postposition de l'adjectif
2.1. Problématique
2.2. Histoire : le romani albanais ancien
2.3. Géographie : le romani des autres régions
2.3.1. Descriptions d'autres variétés
2.3.2. Données roumaines contemporaines
2.3.3. Données roumaines anciennes
2.3.4. Autres données
2.4. Scénario géographique
2.4.1. Dans les Balkans
2.4.2. Hors des Balkans
2.5. Scénario historique
2.5.1. Problème de chronologie des acquisitions
2.5.2. Solution de chronologie
2.5.3. Objections
3. Scénario socio-linguistique
3.1. Types d'résultats
3.2. Types de résultats
3.3. Mes résultats
3.4. Interprétation de nos résultats
3.4.1. DEF A N
3.4.2. DEF N A
3.4.3. DEF N DEF A
4. Scénario théorique
4.1. Pourquoi est-elle cette structure ?
4.1.1. L'article défini supplémentaire
4.1.2. La place de l'adjectif
4.1.3. Les fonctions
4.1.4. L'optionnalité
4.2. La mécanisme de changement syntaxique
4.2.1. La mécanisme à l'œuvre en romani
4.2.2. Le mécanisme à l'œuvre en grec
4.2.3. Le processus de transfert linguistique
5. Réponse à la problématique
Chapitre V : corpus semi-spontané
1. Méthodologie
1.1. Introduction
1.2. Les textes réalisés
1.2.1. Test 1
1.2.2. Test 2
1.2.3. Récit
1.3. Découvrement et codage
2. Résultats : variables linguistiques
2.1. Distribution des syntagmes définis
2.1.1. Test 1
2.1.2. Test 2
2.1.3. Récit
2.1.4. Témoins
2.1.5. Synthèse
2.2. Variable 1 : DÉFINITUDE
2.2.1. Quelle règle ?
2.2.2. Régler l'article défini en référence d'un autre déterminant
2.2.3. Pourquoi ?
2.3. Variable 2 : ORDRE DES CONSTITUANTS
2.3.1. Syntagmes monodéfinis
2.3.2. Syntagmes polydéfinis
2.3.3. Syntagmes introduits par un autre déterminant
2.3.4. Synthèse
2.4. Variable 3 : TYPE D'ADJECTIF
2.4.1. Poids des adjectifs et type syntaxique
2.4.2. Origine et sémantique des adjectifs dans les syntagmes polydéfinis
2.4.3. Origine des adjectifs dans les syntagmes définis
2.4.4. Sémantique des adjectifs dans les syntagmes définis
2.4.5. La combinaison sémantique adjectif – nom
2.4.6. Témoins
2.4.7. Synthèse
2.5. Synthèse des variables linguistiques
3. Résultats : variables sociales
3.1. Introduction
3.1.1. Variables étudiées
3.1.2. Variables étudiées
3.1.3. Le genre
3.1.4. Le système scolaire en Albanie
3.1.5. Le millieu rural
3.1.6. Les variétés
3.2. Variable 1 : GENRE
3.2.1. Participants albanais
3.2.2. Témoins
3.3. Variable 2 : CLASSE D'AGE
3.3.1. Participants albanais
3.3.2. Témoins
3.4. Variable 3 : EDUCATION
3.4.1. Participants albanais
3.4.2. Témoins
3.5. Variable 4 : MILITARISME
3.5.1. Participants albanais
3.5.2. Témoins
3.6. Variable 5 : MOBILITÉ
3.6.1. Participants albanais
3.6.2. Témoins
3.7. Synthèse des variables sociales
3.7.1. Test 1 – test 2
3.7.2. Hétérogénéité des variables
3.7.3. Profil des locuteur-rices
3.7.4. Perspectives
3.7.5. Synthèse
4. Synthèse générale
4.1. Exploration du corpus spontané
4.2. Les informants
4.3. L'usage du romani au quotidien
4.3.1. Dans la famille
4.3.2. Hors de la famille
4.4. Le matériel et les données collectées
4.5. La transcription et la présentation des examples
4.6. Analyse des données
Chapitre IV : corpus spontané
1. Méthodologie
2. Syntagmes sans rupture interne
2.1. Syntagmes polydéfinis
2.1.1. DEF N DEF A
2.1.2. DEF A DEF N
2.1.3. Synthèse
2.2. Syntagmes monodéfinis
2.2.1. DEF A N
2.2.2. DEF N A
2.2.3. Synthèse
2.3. Synthèse générale
2.3.1. Type d'adjectif
2.3.2. Modification d'un type d'adjectif
2.3.3. Place de l'adjectif
2.3.4. Place de l'adjectif et variation dialectale
2.3.5. Pourquoi un deuxième article défini ?
2.3.6. Réponse à la problématique
3. Syntagmes avec rupture interne
3.1. Syntagmes polydéfinis
3.1.1. DEF N DEF A
3.1.2. DEF A/ DEF N
3.1.3. Synthèse
3.2. Syntagmes monodéfinis
3.2.1. DEF A/ N
3.2.2. DEF N/ A
3.2.3. Synthèse
3.3. Synthèse générale
3.3.1. Type d'adjectif
3.3.2. Modification d'un type d'adjectif
3.3.3. Place de l'adjectif et variation dialectale
3.3.4. Bilan de la rupture et de l'insertion
3.3.5. Réponse à la problématique

Table des figures

Figure 1 : Classification conventionnelle des dialectes en linguistique romani selon (Matras 2010 : 56) 33
Figure 2 : Les isoglosses majeures et la grande fracture selon (Matras 2010 : 5) 34
Figure 3 : Les Balkans et la ville de Korçë (Albanie) 38
Figure 4 : Processus permettant de décrire une langue sur la base de l'intuition et des jugements d'acceptabilité 45
Figure 5 : Vue des alentours de la maison de mes hôtes à Korçë 51
Figure 6 : Continuum de la définitude 63
Figure 7 : Processus de grammaticalisation de l'article défini (De Mulder et Carlier 2011 : 524) 67
Figure 8 : Échelle d'individuation (De Mulder et Carlier 2011 : 524) 69
Figure 9 : Ordre d'apparition des adjectifs dans différentes langues selon (Cinque 2010 : 38) 77
Figure 10 : Ordre d'apparition des adjectifs selon Stavrou (1999 : 203) 77
Figure 11 : Ordre des constituants nominaux 77
Figure 12 : Ordre d'apparition des adjectifs en anglais (Laenzlinger 2000 : 59-60) 79
Figure 13 : Lecture intersective *alias* interprétation parallèle du syntagme nominal (Stavrou 1999 : 208) 85
Figure 14 : Lecture intersective *alias* interprétation hiérarchique du syntagme nominal (Stavrou 1999 : 208) 86
Figure 15 : Interprétations intersectives de l'adjectif *petit* 87
Figure 16 : Résumé des propriétés du syntagme adjectival selon sa source syntaxique (Cinque 2010) 88
Figure 17 : Termes exprimant l'opposition *topic/focus* (Nølke 2001 : 89) 91
Figure 18 : Termes exprimant l'opposition entre focus contrastif et focus informatif 95
Figure 19 : Topologie du syntagme nominal (Nølke 2001 : 199, 209) 99
Figure 20 : Schéma résumant les analyses de Cinque (2010) 101
Figure 21 : Résumé de l'ordre des adjectifs dans un syntagme nominal en anglais et en italien (Cinque 2010 : 22) 101
Figure 22 : Fusion de l'article défini avec une préposition 109
Figure 23 : Locutrice de l'exemple (12) 115
Figure 24 : Cuisson du bourek (*marikli*) au feu de bois à Bilisht 115
Figure 25 : Structure sous-jacente d'un syntagme polydéfini grec (Campos et Stavrou 2004 : 157) 165
Figure 26 : Interprétation non-restrictive du syntagme (126)a,b 181
Figure 27 : Interprétation restrictive du syntagme (126)a,c 181
Figure 28 : La Tour Blanche (Thessalonique) 189
Figure 29 : Problématique 200
Figure 30 : Sélection d'un petit frère parmi les frères et sœurs 214
Figure 31 : La brasserie de Korçë 218
Figure 32 : SN polydéfini synthétisant deux SN précédents 220
Figure 33 : Redistribution fonctionnelle des fonctions de l'adjectif épithète en diachronie 250
Figure 34 : Esquisse de réponse à la problématique 262
Figure 35 : Début de la réponse à la problématique 285
Figure 36 : De la langue au corpus, du corpus à la langue (Fagard 2015 : 11) 288
Figure 37 : Arrangement préalable des objets sur la table de A (test 1) 291
Figure 38 : Exécution du test 1 292
Figure 39 : Arrangement libre des objets sur la table de A (test 2) 295
Figure 40 : Structure des SN définis produits par les participants du test 1 300
Figure 41 : Structure des SN définis produits par les participants du test 2 301
Figure 42 : Structure des SN définis produits par les participants des tests 1 et 2 301
Figure 43 : Structure des SN définis produits par les témoins 302

Figure 44 : Répartition de l'ensemble des crayons noirs, verts et rouges 308
Figure 45 : Échelle d'héritage des syntagmes définis 337
Figure 46 : Trois types de préférences syntaxiques en fonction du TYPE D'ADJECTIF 344
Figure 47 : Proportion décroissante d'adjectifs hérités en fonction des types sémantiques 344
Figure 48 : Hiérarchie sémantique portant à l'usage de la polydéfinitude 347
Figure 49 : Structure des SN définis produits par les participants des tests 1 et 2 en fonction du GENRE 365
Figure 50 : Proportion d'adjectifs anté- et postposés produits par les participants des tests 1 et 2 en fonction du GENRE 366
Figure 51 : Structure des SN définis produits par les témoins en fonction du GENRE 367
Figure 52 : Proportion d'adjectifs anté- et postposés produits par les témoins en fonction du GENRE 367
Figure 53 : Structure des SN définis produits par les participants des tests 1 et 2 en fonction de la CLASSE D'AGE 368
Figure 54 : Proportion d'adjectifs anté- et postposés produits par les participants des tests 1 et 2 en fonction de la classe d'âge 369
Figure 55 : Structure des SN définis produits par les témoins en fonction de la classe d'âge 370
Figure 56 : Proportion d'adjectifs anté- et postposés produits par les témoins en fonction de la classe d'âge 370
Figure 57 : Structure des SN définis produits par les participants des tests 1 et 2 en fonction de l'EDUCATION 371
Figure 58 : Proportion d'adjectifs anté- et postposés produits par les participants des tests 1 et 2 en fonction de l'EDUCATION 372
Figure 59 : Structure des SN définis produits par les témoins en fonction de l'EDUCATION 373
Figure 60 : Proportion d'adjectifs anté- et postposés produits par les témoins en fonction de l'EDUCATION 374
Figure 61 : Structure des SN définis produits par les participants des tests 1 et 2 en fonction du MILITANTISME 375
Figure 62 : Proportion d'adjectifs anté- et postposés produits par les participants des tests 1 et 2 en fonction du MILITANTISME 376
Figure 63 : Structure des SN définis produits par les témoins en fonction du MILITANTISME 376
Figure 64 : Proportion d'adjectifs anté- et postposés produits par les témoins en fonction du MILITANTISME 377
Figure 65 : Structure des SN définis produits par les participants des tests 1 et 2 en fonction de la VARIETE 378
Figure 66 : Proportion d'adjectifs anté- et postposés produits par les participants des tests 1 et 2 en fonction de la VARIETE 379
Figure 67 : Structure des SN définis produits par les témoins en fonction de la VARIETE 380
Figure 68 : Proportion d'adjectifs anté- et postposés produits par les témoins en fonction de la VARIETE 381
Figure 69 : Ordre d'impact décroissant des variables sociales 383
Figure 70 : Regroupement des modalités convergentes issues de GENRE, CLASSE D'AGE, EDUCATION et VARIETE dans le pourcentage de production des différents SN définis 385
Figure 71 : Regroupement des modalités convergentes issues de GENRE, CLASSE D'AGE, EDUCATION et VARIETE dans le pourcentage de production d'adjectifs anté- et postposés 385
Figure 72 : Pourcentage de production des différents SN définis par les participants des tests 1 et 2 en fonction de INDIVIDU 389
Figure 73 : Pourcentage de production d'adjectifs anté- et postposés par les participants des tests 1 et 2 en fonction de INDIVIDU 391
Figure 74 : Suite de la réponse à la problématique 405
Figure 75 : Zone d'extension de la polydéfinitude en romani d'Albanie en 1931 411
Figure 76 : Zone d'extension de la polydéfinitude dans les Balkans 424
Figure 77 : Aires des isoglosses de la mono- et de la polydéfinitude en Europe 435
Figure 78 : Interprétation et ordre des constituants du SN « les lois importantes » en roumain 454
Figure 79 : Scénario historique et géographique pour l'acquisition de la polydéfintiude et de la postposition de l'adjectif dans les différentes variétés de romani 460
Figure 80 : Les quatre configurations de changement linguistique à échelle individuelle et collective (Sankoff 2006 : 111, d'après Labov 1994) 463

Figure 81 : Les cinq configurations possibles de changement linguistique à échelle individuelle et collective (Sankoff et Blondeau 2007 : 563) 464
Figure 82 : Continuum de focalisation sur l'adjectif épithète dans le SN romani 470
Figure 83 : Topologie du syntagme nominal en romani d'Albanie 474
Figure 84 : Les principaux types de transfert linguistique induits par le contact (Heine 2008 : 37) 479
Figure 85 : Le processus de *pattern replication* et de *pivot-matching* (Matras 2009 : 242) 481
Figure 86 : Synthèse du scénario diachronique pour l'acquisition de la polydéfinitude et de la postposition de l'adjectif en romani d'Albanie (Tirard 2017 : 154) 483
Figure 87 : Fin de la réponse à la problématique 504
Figure 88 : *The Patrilineal Kinship Structure of One Roma Fis* (De Soto, Beddies et Gedeshi 2005 : 222) 540
Figure 89 : *Structure of Arben's Fis in Korça* (De Soto, Beddies et Gedeshi 2005 : 22) 541
Figure 90 : *Function of Arben's Fis in Korça* (De Soto, Beddies et Gedeshi 2005 : 23) 541
Figure 91 : Codage du profil des locuteurs à chaque occurrence du corpus spontané 620
Figure 92 : Codage de l'analyse linguistique pour chaque occurrence du corpus spontané 621
Figure 93 : Structure des SN définis et indéfinis comprenant un adjectif produit par les participants des tests 1 et 2 . 627

Table des tableaux

Tableau 1 : Estimation de la population rom et égyptienne en Albanie entre 1991 et 2005 (De Soto, Beddies et Gedeshi 2005 : xxiv) 7
Tableau 2 : Organigramme du système scolaire albanais 9
Tableau 3 : Quelques données sociologiques comparatives (Sula Raxhimi 2015 : 236) 12
Tableau 4 : Endonymes et exonymes des Roms, Égyptiens et Albanais 19
Tableau 5 : Taux d'alphabétisation des Roms par groupes (De Soto, Beddies et Gedeshi 2005 : 189) 26
Tableau 6 : Schéma du répertoire linguistique des informateurs 36
Tableau 7 : Hypothèse d'histoire du contact de langues des variétés de romani d'Albanie 43
Tableau 8 : Résumé de la comparaison entre les propriétés interprétatives de l'adjectif anténominal et postnominal en anglais et italien (Cinque 2010 : 16-17) 75
Tableau 9 : Résumé des propriétés sémantiques et syntaxiques des deux sources de l'adjectif (Cinque 2010 : 27, 33) 76
Tableau 10 : Synthèse des catégories nominales 106
Tableau 11 : Morphologie de l'article défini dans les groupes de dialectes du romani (Matras 2002 : 97) 109
Tableau 12 : Morphologie de l'article défini des trois principales variétés d'Albanie 109
Tableau 13 : Fonctions de l'article défini selon (Boretzky 2000b : 44-54) 112
Tableau 14 : Formes du démonstratif en mečkar et arli 122
Tableau 15 : Formes du possessif dans les variétés d'Albanie 126
Tableau 16 : Les différents noms donnés aux types d'apposition et à leurs sous-parties 171
Tableau 17 : Synthèse des différents types d'appositions en grec moderne 177
Tableau 18 : Répartition par genre et âge des locuteurs enregistrés dans le corpus spontané 202
Tableau 19 : Nombre d'enregistrements effectués en Albanie 207
Tableau 20 : *Tags of the Extended Annotation Scheme for Information Structure* (Götze et al. 2007 : 148-150) 209
Tableau 21 : Huit structurations possibles du syntagme nominal avec adjectif épithète 210
Tableau 22 : Analyse informationnelle et sémantique des syntagmes def n def a à effet de sélection 217
Tableau 23 : Analyse informationnelle et sémantique des syntagmes def n def a à effet d'emphase 219
Tableau 24 : Analyse informationnelle et sémantique des syntagmes def n def a à effet hybride 222
Tableau 25 : Analyse informationnelle et sémantique des syntagmes def a def n 223
Tableau 26 : Le focus en syntagme polydéfini 226
Tableau 27 : Analyse informationnelle et sémantique des syntagmes def a n 230
Tableau 28 : Analyse informationnelle et sémantique des syntagmes def n a 234
Tableau 29 : Focus en syntagme monodéfini 235
Tableau 30 : Poids et place de l'adjectif : quelques exemples du corpus spontané 239
Tableau 31 : Poids et place du nom : quelques exemples du corpus spontané 240
Tableau 32 : Modification directe et indirecte en syntagme monodéfini 241
Tableau 33 : Modification directe et indirecte en syntagme polydéfini 241
Tableau 34 : Profil des locuteurs des syntagmes polydéfinis 243
Tableau 35 : Profil des locuteurs des syntagmes monodéfinis 244
Tableau 36 : Complément au Tableau 117, p. 547 248
Tableau 37 : Complément au Tableau 29, p. 234 249
Tableau 38 : Usages de l'article défini dans les syntagmes définis 252
Tableau 39 : Degré de restrictivité des syntagmes définis 254
Tableau 40 : Quatre structurations possibles du syntagme nominal avec adjectif épithète 261
Tableau 41 : Pause et insertion dans le corpus d'appositions lâches 263

Tableau 42 : Analyse informationnelle et sémantique des syntagmes def n/ def a sans insertion 265
Tableau 43 : Analyse informationnelle et sémantique des syntagmes def n/ def a avec insertion 267
Tableau 44 : Analyse informationnelle et sémantique des syntagmes à adjectif substantivé 268
Tableau 45 : Le focus en syntagme polydéfini avec rupture interne 270
Tableau 46 : Analyse informationnelle et sémantique des syntagmes def a/ n 271
Tableau 47 : Analyse informationnelle et sémantique des syntagmes def n/ a 272
Tableau 48 : Ordre des constituants et focus en syntagme monodéfini avec rupture interne 273
Tableau 49 : Poids et place de l'adjectif avec rupture interne : quelques exemples du corpus spontané 274
Tableau 50 : Poids et place du nom avec rupture interne : quelques exemples du corpus spontané 274
Tableau 51 : Modification directe et indirecte en syntagme polydéfini avec rupture interne 275
Tableau 52 : Modification directe et indirecte en syntagme monodéfini avec rupture interne 275
Tableau 53 : Profil des locuteurs des syntagmes polydéfinis avec rupture intene 275
Tableau 54 : Profil des locuteurs ayant prononcé les syntagmes à adjectif substantivé 276
Tableau 55 : Profil des locuteurs des syntagmes monodéfinis avec rupture interne ou ellipse 276
Tableau 56 : Degré de restrictivité des syntagmes définis avec rupture interne 277
Tableau 57 : Liste et caractéristiques des objets utilisés lors du test 1 290
Tableau 58 : Nombre de locuteurs ayant participé au test 1 en fonction du GENRE 292
Tableau 59 : Nombre de locuteurs ayant participé au test 1 en fonction de la CLASSE D'AGE 293
Tableau 60 : Liste et caratéristiques des objets utilisés lors du test 2 295
Tableau 61 : Nombre de locuteurs ayant participé au test 2 en fonction du GENRE 296
Tableau 62 : Nombre de locuteurs ayant participé au test 2 en fonction de la CLASSE D'AGE 296
Tableau 63 : Déterminants rédupliqués dans les tests 1 et 2 303
Tableau 64 : Déterminants rédupliqués par les témoins 303
Tableau 65 : Position de l'adjectif dans les syntagmes monodéfinis produits par les participants des tests 1 et 2 306
Tableau 66 : Position de l'adjectif dans les syntagmes monodéfinis produits par les témoins 309
Tableau 67 : Position de l'adjectif dans les syntagmes polydéfinis produits par les participants des tests 1 et 2 311
Tableau 68 : Position de l'adjectif dans les autres syntagmes produits par les participants des tests 1 et 2 325
Tableau 69 : Position de l'adjectif dans les autres syntagmes produits par les témoins 326
Tableau 70 : Economie du syntagme nominal dans les tests 1 et 2 : morphologie et syntaxe 328
Tableau 71 : Typologie de l'emprunt adjectival chez Matras (2009) 329
Tableau 72 : Poids et place de l'adjectif : quelques exemples du corpus semi-spontané 331
Tableau 73 : Poids et place du nom : quelques exemples du corpus semi-spontané 332
Tableau 74 : Morphologie des adjectifs empruntés dans les syntagmes polydéfinis des tests 1 et 2 (témoins exclus) 333
Tableau 75 : Origine et type sémantique des adjectifs des syntagmes polydéfinis des tests 1 et 2 (témoins exclus) ... 334
Tableau 76 : Synthèse de l'origine des adjectifs en fonction de leur sémantique des syntagmes polydéfinis des tests 1 et 2 (témoins exclus) 335
Tableau 77 : Synthèse de l'origine des adjectifs dans tous les SN définis des tests 1 et 2 (témoins exclus) 336
Tableau 78 : Préférences structurelles des adjectifs en fonction de leur origine dans les SN définis des tests 1 et 2 (témoins exclus) 337
Tableau 79 : Analyse des occurrences d'adjectifs empruntés dans les syntagmes def a n 338
Tableau 80 : Morphologie des adjectifs empruntés dans les SN monodéfinis et complexes des tests 1 et 2 (témoins exclus) 339
Tableau 81 : Analyse des occurrences d'adjectifs empruntés dans les syntagmes def n a 340
Tableau 82 : Préférences syntaxiques des adjectifs en fonction de leur sémantique dans les SN définis des tests 1 et 2 (témoins exclus) 343
Tableau 83 : Synthèse de l'origine des adjectifs en fonction de leur sémantique dans les SN définis des tests 1 et 2 (témoins exclus) 344
Tableau 84 : Type d'interprétation des adjectifs des syntagmes polydéfinis des tests 1 et 2 (témoins exclus) 346
Tableau 85 : Analyse des occurrences d'adjectifs empruntés antéposés chez les témoins 348

Tableau 86 : Synthèse de l'origine des adjectifs dans tous les SN définis des témoins 349
Tableau 87 : Analyse des occurrences d'adjectifs empruntés postposés chez les témoins 350
Tableau 88 : Morphologie des adjectifs empruntés dans les SN définis produits par les témoins 350
Tableau 89 : Modèles syntaxiques des adjectifs empruntés par les participants albanais (hors témoins) 351
Tableau 90 : Modèles morphologiques des adjectifs empruntés par les participants albanais 353
Tableau 91 : Récapitulatif des variables sociales, des modalités, et de l'échantillon étudié 359
Tableau 92 : L'engagement militant de l'ensemble des participants aux tests 1 et 2 363
Tableau 93 : Profil de locuteur le plus probable pour chaque syntaxe nominale 383
Tableau 94 : Contes collecté par Stuart Mann en 1931 409
Tableau 95 : Relevé des SN définis avec adjectif épithète dans les contes collectés par Stuart Mann 410
Tableau 96 : Variétés du corpus RMS comportant de la polydéfinitude 412
Tableau 97 : Syntagmes polydéfinis dans le corpus RMS 413
Tableau 98 : Exemples de syntagmes polydéfinis dans les textes collectés par Barbu Constantinescu 417
Tableau 99 : Présence de la mono- et de la polydéfinitude dans les autres langues des Balkans 426
Tableau 100 : Hypothèse de présence de la mono- et de la polydéfinitude en Europe 436
Tableau 101 : Présence de la mono- et de la polydéfinitude dans les principales langues d'Europe 437
Tableau 102 : Le marquage de la définitude dans les langues germaniques (Schoorlemmer 2012 : 110) 439
Tableau 103 : Variation morphologique de la définitude dans les langues germaniques (Schoorlemmer 2012 : 153) . 439
Tableau 104 : Scénario d'une rétention (hypothèse I) 445
Tableau 105 : Scénario d'un ancien emprunt (hypothèse II) 445
Tableau 106 : Exemples de syntagmes avec adjectif antéposé dans les textes collectés par Barbu Constantinescu 449
Tableau 107 : Présence de la postposition de l'adjectif dans les autres langues des Balkans 452
Tableau 108 : Hypothèse de présence de postposition de l'adjectif en Europe 455
Tableau 109 : Présence de la postposition de l'adjectif dans les principales langues d'Europe 456
Tableau 110 : Scénario d'un ancien emprunt (hypothèse II') 458
Tableau 111 : Les différents types de redoublement syntaxique Barbiers (2013 : 9-10) 478
Tableau 112 : Chronologie de l'acquisition de l'article défini, de la polydéfinitude et de la postposition de l'adjectif dans les variétés balkaniques 500
Tableau 113 : Typologie du marquage multiple de la définitude dans les langues d'Europe 502
Tableau 114 : Profil des locuteurs ayant prononcé les exemples spontanés 543
Tableau 115 : Analyse des sèmes de *čhavo* et de *rom* 544
Tableau 116 : Synthèse des analyses présentées pour les exemples polydéfinis 545
Tableau 117 : Synthèse des analyses présentées pour les exemples monodéfinis 546
Tableau 118 : Synthèse des analyses présentées pour les exemples polydéfinis avec rupture interne 547
Tableau 119 : Synthèse des analyses présentées pour les syntagmes à adjectif substantivé 547
Tableau 120 : Synthèse des analyses présentées pour les exemples monodéfinis avec rupture interne 548
Tableau 121 : Les couples de participants du test 1 610
Tableau 122 : Les couples de participants du test 2 610
Tableau 123 : Profil des locuteurs ayant participé aux tests 1 et 2 613
Tableau 124 : Codage des données 616
Tableau 125 : Les structures et déterminants du corpus : avec quels adjectifs les trouve-t-on ? 622
Tableau 126 : Les adjectifs du corpus : avec quelles structures ou déterminants les trouve-t-on ? 625
Tableau 127 : Structure des SN définis comprenant un adjectif produit par les participants des tests 1 et 2 627
Tableau 128 : Structure des SN définis comprenant un adjectif produit par les témoins 627
Tableau 129 : Structure des SN définis et indéfinis comprenant un adjectif produit par les participants des tests 1 et 2 627
Tableau 130 : Structure des SN définis produits par les participants des tests 1 et 2 en fonction du GENRE 628
Tableau 131 : Proportion d'adjectifs anté- et postposés produits par les participants des tests 1 et 2 en fonction du GENRE 628

Tableau 132 : Structure des SN définis produits par les témoins en fonction du GENRE 628
Tableau 133 : Proportion d'adjectifs anté- et postposés produits par les témoins en fonction du GENRE 628
Tableau 134 : Structure des SN définis produits par les participants des tests 1 et 2 en fonction de la CLASSE D'AGE 629
Tableau 135 : Proportion d'adjectifs anté- et postposés produits par les participants des tests 1 et 2 en fonction de la CLASSE D'AGE 629
Tableau 136 : Structure des SN définis produits par les témoins en fonction de la CLASSE D'AGE 629
Tableau 137 : Proportion d'adjectifs anté- et postposés produits par les témoins en fonction de la CLASSE D'AGE 629
Tableau 138 : Structure des SN définis produits par les participants des tests 1 et 2 en fonction de l'EDUCATION 630
Tableau 139 : Proportion d'adjectifs anté- et postposés produits par les participants des tests 1 et 2 en fonction de l'EDUCATION 630
Tableau 140 : Structure des SN définis produits par les témoins en fonction de l'EDUCATION 630
Tableau 141 : Proportion d'adjectifs anté- et postposés produits par les témoins en fonction de l'EDUCATION 630
Tableau 142 : Structure des SN définis produits par les participants des tests 1 et 2 en fonction du MILITANTISME 631
Tableau 143 : Proportion d'adjectifs anté- et postposés produits par les participants des tests 1 et 2 en fonction du MILITANTISME 631
Tableau 144 : Structure des SN définis produits par les témoins en fonction du MILITANTISME 631
Tableau 145 : Proportion d'adjectifs anté- et postposés produits par les témoins en fonction du MILITANTISME 631
Tableau 146 : Structure des SN définis produits par les participants des tests 1 et 2 en fonction de la VARIETE 631
Tableau 147 : Proportion d'adjectifs anté- et postposés produits par les participants des tests 1 et 2 en fonction de la VARIETE 632
Tableau 148 : Structure des SN définis produits par les témoins en fonction de la VARIETE 632
Tableau 149 : Proportion d'adjectifs anté- et postposés produits par les témoins en fonction de la VARIETE 632
Tableau 150 : Regroupement des modalités convergentes issues de GENRE, CLASSE D'AGE, EDUCATION et VARIETE dans le pourcentage de production des différents SN définis 633
Tableau 151 : Regroupement des modalités convergentes issues de GENRE, CLASSE D'AGE, EDUCATION et VARIETE dans le pourcentage de production d'adjectifs anté- et postposés 633
Tableau 152 : Pourcentage de production des différents SN définis par les participants des tests 1 et 2 en fonction de INDIVIDU 634
Tableau 153 : Pourcentage de production d'adjectifs anté- et postposés par les participants des tests 1 et 2 en fonction de INDIVIDU 635
Tableau 154 : Profil des locuteurs ayant prononcé les exemples spontanés du chapitre VI 636

Table des abréviations

Partie du discours	Abréviation	Partie du discours	Abréviation
première personne	1	locatif	LOC
deuxième personne	2	masculin	M
troisième personne	3	nom	N
adjectif	A	nom propre	NP
ablatif	ABL	négation	NEG
adverbe	ADV	numéral	NUM
auxiliaire	AUX	cas oblique	OBL
comparatif	COMP	participe	PART
complémenteur	COMPL	passif	PASS
datif	DAT	pluriel	PL
article défini	DEF	possessif	POSS
démonstratif	DEM	préposition	PREP
cas direct	DIR	présent	PRES
féminin	F	prétérit	PRET
focalisateur	FOC	participe	PTCP
futur	FUT	quantifieur	QUANT
génitif	GEN	réfléchi	REFL
impératif	IMP	pronom/adverbe relatif	REL
imparfait	IMPF	singulier	SG
article indéfini	IND	syntagme génitival	SGEN
instrumental	INSTR	syntagme nominal	SN
interrogatif	INT	proposition relative	SREL

Introduction

Le présent travail porte sur les syntagmes nominaux comportant un adjectif épithète articulé en romani d'Albanie. Le romani est une langue indo-européenne de la branche indo-iranienne. Elle est parlée sur plusieurs continents, notamment européen et américain, par environ onze millions de personnes selon le rapport du Conseil de l'Europe (2012), qui propose une estimation variant de huit à quinze millions. Cette langue comprend de très nombreuses variétés locales en raison de la dispersion des Roms sur plusieurs continents il y a environ mille ans, et de leur vie en petites communautés partout minoritaires. La situation historique et sociolinguistique du romani en fait une langue en contact permanent depuis son exode de l'Inde (Matras 2002 : 191). Ainsi, il n'existe aucun locuteur monolingue du romani, à l'exception peut-être des très jeunes enfants : le plurilinguisme est constitutif de la vie des Roms qui le parlent. On comprend que le contact de langues soit un élément déterminant de son étude et qu'il faille pour cela prendre en compte un nombre important de langues, différentes selon les époques et les régions.

Les syntagmes qui font l'objet de cette étude s'opposent à ceux qui ne comportent pas d'adjectif, à ceux où l'adjectif n'est pas épithète du nom, et surtout à ceux où il n'est pas accompagné d'un article défini propre (syntagmes que nous avons appelés « monodéfinis »). Il s'agit de marquage multiple de la définitude au sein d'un syntagme nominal unique, qui ne présente ni pause à l'oral ni virgule à l'écrit. Ce phénomène n'a jamais été étudié pour le romani, mais on peut le rapprocher de structures similaires présentes dans d'autres langues telles que le grec, le roumain ou le suédois. Pour le grec, Lekakou et Szendrői (2012 : 108) le définissent comme des « *instances of an adjective modifying a noun where the noun and the adjective are each accompanied by their own determiner* ». Il porte des noms différents dans la littérature scientifique :

- syntagmes polydéfinis – *polydefinite NPs / polydefinites* (Kolliakou 2004, Lekakou et Szendrői 2007)
- expansion/diffusion du déterminant – *determiner spreading* (Androutsopoulou 1995)
- déterminants multiples – *multiple determiners* (Alexiadou et Wilder 1998, Alexiadou 2014)
- construction à double article adjectival – *double definite adjectival article construction* (Cornilescu et Nicolae 2012)
- double articulation – *double articulation* (Plank 2003b)
- double détermination – *double determination* (Börjars 1994)
- double définition – *double definition* (Börjars 1994)
- surdétermination – *overdetermination, overbestemthet* (Boretzky 2000b, Plank 2003b)

Ce phénomène n'est pas systématique mais fait l'objet d'une variation libre en romani d'Albanie, de même que la place de l'adjectif. Lors de nos enquêtes de terrain, nous avons par exemple relevé les syntagmes suivants[1] :

a.	*i*	*bari*	*mačka*
	la	grande	chatte

[1] Une chatte venait de mettre bas dans le cellier de la maison où j'ai effectué mon séjour de terrain en 2014.

b.	*i*	*mačka*	*bari*	
	la	chatte	grande	
c.	*i*	*mačka*	*i*	*bari*
	la	chatte	la	grande

Ces syntagmes se traduisent de la même façon en français, « la grande chatte ». Ils ne semblent pas présenter de différence sémantique – et n'en présentent pas, au dire des locuteurs. Ce phénomène se caractérise donc par l'absence apparente d'ajout de sens, c'est pourquoi les locuteurs ont difficilement conscience d'employer cette structure. Si le syntagme est polydéfini, son sens reste monodéfini.

Le fait d'ajouter un article défini dans un syntagme le rend plus complexe sur le plan grammatical, mais aussi plus long à prononcer sur le plan communicatif. Le redoublement de l'article défini semble une pure redondance grammaticale : à quoi sert-elle donc ? Quel est son rôle dans la langue ? Elle semble aller contre les principes de simplicité et d'économie, contre la « loi du moindre effort ». Or nous pensons que les locuteurs ne font rien en vain et que tout, dans la langue, a une utilité.

Nous examinerons les conditions dans lesquelles s'emploie cette construction. Quel est le rapport entre marquage double/multiple et marquage simple de la définitude ? Nous devrons pour cela examiner la structure du syntagme nominal, le rôle du ou des déterminants et celui de l'adjectif. On observe empiriquement que cette construction présente le plus souvent un adjectif postposé au nom : quel est le lien entre marquage double/multiple de la définitude et ordre des constituants nominaux ? Ce phénomène fait-il l'objet d'une variation réellement « libre » ?

Cette étude vise à le décrire et à l'expliquer dans les variétés albanaises du romani et si possible pour les autres variétés. Existe-t-il dans toutes ? Quelles différences observe-t-on avec les structures similaires décrites pour les autres langues ? Nous pourrons notamment comparer la structure du romani albanais à celle du grec, très étudiée depuis l'article fondateur d'Androutsopoulou (1995).

Une question essentielle constituera le fil rouge de ce travail : *pourquoi* ? Quelle raison motive l'emploi d'une telle structure alors qu'elle n'est pas obligatoire et ne semble pas modifier le sens du syntagme ? Nous chercherons à comprendre *comment, quand, à quelle fréquence, dans quel(s) contexte(s)* sont employées les constructions polydéfinies, mais aussi *par qui, par quelles variétés, depuis quand* et *où*. Nous répondrons à un ensemble de questions d'ordre morphologique, syntaxique, sémantique, fonctionnel, socio-linguistique, typologique et diachronique.

Pour effectuer ce travail, j'ai réalisé plusieurs missions de terrain à Korçë en 2013 et 2014, au cours desquelles j'ai communiqué exclusivement en romani avec mes interlocuteurs. Korçë se trouve dans une région d'Albanie particulièrement intéressante en ce qui concerne le contact de langues, puisque située aux confins de la Grèce et de la Macédoine du Nord. J'y ai collecté un corpus de discours spontané ainsi qu'un corpus semi-spontané issu d'une expérience réalisée en 2014. J'ai donc choisi de combiner des analyses qualitatives et quantitatives systématiques. En effet, le contenu informatif et pragmatique des données acquises en milieu écologique est hautement variable, contingent et imprévisible. Cela rend les données non systématiques et difficilement comparables. Tout au plus peut-on tenter d'influencer le cours et la thématique du discours durant la conversation, ou bien de poser des questions lors d'entretiens semi-dirigés. Au contraire, les données de tests

linguistiques permettent un traitement statistique de l'ordre des constituants et une mesure de la fréquence d'apparition des structures. Mais une telle étude quantitative ne vaut que pour le seul contexte de l'expérience, jugé au préalable favorable à l'apparition de la structure recherchée. Or rien ne nous indique qu'il n'existe pas d'autres contextes, que seuls peuvent dégager l'étude qualitative du discours spontané. C'est pourquoi ces deux types de données et d'analyses sont complémentaires.

Ce travail comporte six parties.

- La première partie présente les Roms d'Albanie et leur langue ; elle explique la méthodologie employée pour mener à bien le présent travail.
- La deuxième constitue une introduction théorique aux facteurs linguistiques qui influent sur les syntagmes polydéfinis : la définitude et l'article défini, l'adjectif, la structure de l'information et le focus.
- La troisième constitue l'état de l'art de la structure polydéfinie. Elle présente le syntagme nominal romani, les quelques mentions de cette structure dans les descriptions du romani ainsi que les études consacrées à la structure similaire en grec.
- La quatrième partie présente les données qui forment le corpus « spontané ». Nous les classons en fonction des types de syntagmes et de la présence ou non d'insertion lexicale ou de pause interne, et les analysons d'une manière qualitative.
- La cinquième présente les données qui forment le corpus « semi-spontané ». Nous les analysons de manière quantitative en distinguant plusieurs variables internes (linguistiques) et externes (socio-linguistiques) à la langue.
- Enfin, la sixième partie rassemble les conclusions des analyses et les replace dans le contexte de la recherche sur la polydéfinitude en général et sur le romani en particulier. Elle explique quel est l'intérêt, pour les locuteurs, d'employer cette structure, et quel rôle y jouent l'article défini « supplémentaire » et la place de l'adjectif. Nous élaborons un scénario socio-linguistique et un scénario diachronique afin d'expliquer l'usage de la polydéfinitude en romani d'Albanie ainsi que dans les autres variétés balkaniques de la langue. Enfin, nous tentons de dresser une typologie des syntagmes polydéfinis dans les langues d'Europe afin d'y situer ceux du romani.

Cette étude s'achève sur une conclusion générale des résultats obtenus et sur une mise en perspective. Suit un ensemble d'annexes qui comporte, entre autres, la bibliographie, les métadonnées des enregistrements et des informateurs sur le terrain, le corpus spontané transcrit et traduit en français, le classement des adjectifs du corpus semi-spontané par structure et par déterminant, et les résultats chiffrés de l'analyse du corpus semi-spontané sous forme de tableaux et de diagrammes.

Chapitre I : Les Roms d'Albanie et leur langue

Ce chapitre constitue une introduction au peuple et à la langue qui font l'objet notre étude. Il présente ensuite la méthodologie employée pour mener à bien le travail.

1. Les Roms d'Albanie

1.1. Éléments démographiques

1.1.1. Taille de la population

Les Roms sont présents sur le territoire albanophone[1] depuis plus de 600 ans (Kolsti 1991 : 4), soit avant l'instauration de l'Empire ottoman. Durant cette période, les Roms semblent avoir vécu en cohabitation avec les autres groupes ethniques, nombreux dans cet empire multi-ethnique. On a peu d'informations sur leur histoire, sinon en consultant les notes et descriptions du voyageur Evliya Çelebi tout au long de sa vie en dix volumes (Çelebi 1630). Les registres fiscaux de l'empire fournissent également nombre d'informations sur le nombre de familles et leur localisation. Le registre intégral de la population rom de Bulgarie, Macédoine, Serbie, Albanie et Grèce pour l'année 1523 a été résumé et publié par l'historien macédonien Stojanovski (Stojanovski 1989 : 127-180). C'est la première mention des Roms albanais dans les textes anciens.

Il est difficile de savoir quelle était la taille de cette population. On peut trouver les estimations suivantes dans les études :

- 965 individus lors du recensement ottoman de 1523, répartis en plus de 300 foyers[2] et dont 32 veuves (Stojanovski 1989 : 127-180)[3]. Notons que 861 individus sont recensés comme chrétiens et 72 individus comme musulmans[4].
- 100 familles à Tepelen (70 musulmanes et 30 chrétiennes), soit près d'un demi-millier de personnes vers 1909 selon (Durham 1905 : 288),
- 60 à 80 000 en 1918 en Macédoine slave et grecque, selon l'estimation de l'ethnologue (Weigand 1924 : 86),

[1] L'État albanais indépendant n'a été proclamé que le 28 novembre 1912, et n'a commencé à fonctionner réellement qu'après la Première Guerre Mondiale.
[2] Marcel Courthiade compte quant à lui 1270 individus répartis en 374 foyers selon (Koinova 2000 : 7).
[3] Nous tenons à remercier Bernard Lory pour la traduction du résumé original en macédonien.
[4] Dont quarante-sept résident dans le district de Priština, dix dans celui de Prizren et dix dans celui d'Ohrid.

- 20 000 dans les années 1930 (Hasluck 1938a : 50),
- 62 000 en 1980 (*La documentation française* citée par Koinova 2000 : 7),
- 85 000 en 1984 (Marcel Courthiade, dans sa traduction de (Hoxha 1984), sur une population totale de trois millions d'habitants,
- 95 000 entre 1991 et 1994 (De Soto, Beddies et Gedeshi 2005 : xx), soit 2 % de la population totale de 3 421 000 personnes,
- 100 000 selon la plupart des études mentionnées par (De Soto, Beddies et Gedeshi 2005 : xxiv), et qui datent toutes de 1991 à 1997.
- 90 000, dont 95 % connaissent la langue selon (Bakker 2001 : 299) – pourcentage le plus élevé d'Europe,
- 100 000 selon (Gurbetovski [Courthiade] 2010),
- 8 301 à 120 000 dans les années 2010,
- entre 80 000 et 150 000 en 2012 selon le Ad Hoc Committee of Experts on Roma Issues (2012) du Conseil de l'Europe, avec une moyenne de 115 000, soit 3,59 % de la population albanaise (3 204 284 personnes en 2010 selon la Banque Mondiale).

Il est fort possible que, pour des raisons religieuses, la taille de la population rom albanaise ait augmenté dans la deuxième moitié du XIXe siècle (Koinova 2000 : 5). Les quatre vilayets albanais[1] auraient en effet servi de refuge aux Roms musulmans des régions récemment reconquises par les Chrétiens et devenues indépendantes… et d'autre part aux Roms chrétiens fuyant l'esclavage dans les principautés roumaines.[2] L'histoire se répéta en 1924 lors de l'« échange de population » entre la Grèce et la Turquie : les Roms musulmans du nord-ouest (voire du nord-est) de la Grèce préférèrent passer la frontière dans le pays voisin plutôt que d'émigrer dans la lointaine Turquie.

Le grand écart des estimations les plus récentes s'explique par la comptabilisation conjointe ou non des Roms et des Égyptiens[3], et par l'absence de la minorité rom dans les recensements jusqu'à 2011 – contrairement aux minorités grecque et macédonienne, qui figurent au recensement depuis 1945, et serbo-monténégrine depuis 1950. Les Roms se faisaient donc recenser comme Albanais ou comme minoritaires grecs, macédoniens, serbo-monténégrins. Une catégorie « Autres » a vu le jour en 1979 et permis aux Roms, Égyptiens et Aroumains/Vlax de s'y faire recenser. Des catégories séparées pour les Roms, Égyptiens, Aroumains/Vlax et Monténégrins sont apparues en 2011 – la catégorie serbe disparaissant au passage. Selon ce recensement, seules 8 301 personnes se seraient déclarées Roms

[1] Les vilayets albanais englobent, à la fin de la période ottomane, le Kosovo, la Macédoine slave et l'Epire, soit Shkodër, Skopje, Ioaninna et Monastir/Manastir/Bitola.

[2] Pour une esquisse de l'histoire des Roms en Albanie sous l'Empire ottoman et au XXème siècle, cf. Sula Raxhimi (2015 : section 2.2) – qui déplore le « peu d'écrits – scientifiquement solides – sur les Roms en Albanie » (Sula Raxhimi 2015 : 8). Il existe néanmoins un certain nombre de descriptions historiques au XIXème siècle, et on peut trouver des sources plus anciennes encore notamment (Stojanovski 1989).

[3] Il s'agit d'une minorité ethnique albanaise que nous présentons en section 1.4, p. 25.

sur une population totale de 2 800 138 Albanais, soit 0,30 % de la population totale. On peut penser que le reste des Roms a continué à se faire recenser dans la minorité « autre » (2 644 personnes), dans la majorité albanaise (2 312 356 personnes, 82,58 % de la population totale) … voire dans la catégorie « préfère ne pas répondre » (390 938 personnes, 13,96 % de la population !). Le recensement démolinguistique n'est pas plus éclairant : seule 4 025 personnes ont déclaré le romani pour langue maternelle, soit moins de la moitié des Roms officiellement déclarés. Le Conseil de l'Europe ne considère pas le recensement fiable et l'a sévèrement critiqué (Advisory Committee on the Framework Convention for the Protection of National Minorities 2012).

Si l'on fait la moyenne de ces estimations, en prenant soin de différencier Roms et Égyptiens, ce que font De Soto, Beddies et Gedeshi (2005) et Gurbetovski [Courthiade] (2010), mais non le Conseil de l'Europe, on peut estimer que la population rom *stricto sensu* atteint peut-être 110 à 120 000 personnes en 2018. Son taux de croissance démographique est reconnu comme bien plus important que celui des Non-Roms (Koinova 2000). On peut noter que les nombreuses études sociologiques commandées par les organisations internationales, gouvernementales ou non, n'ont pas réussi à mieux mesurer la taille de la population rom d'Albanie, ni sa répartition au sein du pays.

1.1.2. Répartition géographique de la population

Le Tableau 1 présente la répartition des Roms selon De Soto, Beddies et Gedeshi (2005 : xxiv) :

Tableau 1 : Estimation de la population rom et égyptienne en Albanie entre 1991 et 2005 (De Soto, Beddies et Gedeshi 2005 : xxiv)

Table 2. Estimated Population of Albanian Roma and Egyptians, by District

	Local Government		Roma Association Amaro Drom	Poulton[a*]	Bugajski[b*]	State Department[c*]	Brunër[d*]	Egyptian Association Vëllazërimi
Regions	Roma	Egyptian	Roma	Roma	Roma	Roma	Roma	Egyptian
Shkodra	50	4,500	600–1,000	n/a	n/a	n/a	n/a	n/a
Kruja	n/a	n/a	1,500–2,000	n/a	n/a	n/a	n/a	n/a
Tirana	n/a	n/a	10–12,000	n/a	n/a	n/a	n/a	n/a
Elbasan	2,622	8318	9–10,000	n/a	n/a	n/a	n/a	10,000
Durres	n/a	n/a	3–4,000	n/a	n/a	n/a	n/a	20,000
Korça	2,383	n/a	10–11,000	n/a	n/a	n/a	n/a	28,000
Fier	5,323	n/a	17–18,000	n/a	n/a	n/a	n/a	400
Vlora	1,036	3,000	2,000–2,500	n/a	n/a	n/a	n/a	30,000
Berat	n/a	n/a	3–4,000	n/a	n/a	n/a	n/a	10,000
Gjirokastra	1,200	5,000	2,000	n/a	n/a	n/a	n/a	5,000
Delvina	285	2,500	2,000	n/a	n/a	n/a	n/a	1,100
Total	n/a	n/a	120,000–150,000	10,000	10,000	100,000	60,000	Over 200,000

Notes: The official statistics, provided in all the population censuses conducted in Albania since the World War II (2002), do not give separate figures about the Roma and Egyptian communities. These are included in the total Albanian population figures. Meanwhile, in some districts, local governments have made their own estimates of the Roma and Egyptian population. In addition, the Roma and Egyptian associations have made their estimates of these populations in certain districts as well as nationally. Several non-Albanian authors have also attempted to approximate the numbers of these two communities. The striking differences between these estimates show that these figures must be taken with a grain of salt.
a. Poulton, H (1991). "The Balkans: Minorities and States in Conflict." London: Minority Rights Group.
b. Bugajski, J (1994). "Ethnic Politics in Eastern Europe." Armonk, NY: M. E. Sharp.
c. U.S. Department of State (1995). "Country Reports on Human Rights Practices for 1994." Washington, DC: U.S. Government Printing Office, February.
d. Brunër, G (1995). "Nemzetisegi kerdes es kisebbsegi konfliktusor Kelet-Europaban." Budapest: Teleki Laszl Alapitvany.
Source: ERRC (1997). "No Record of the Case: Roma in Albania," *Country Report Series* No. 5, June.

Les Roms sont bien répartis sur l'ensemble du territoire, ce qui est confirmé par Koinova (2000 : 7-8) et European Roma Rights Center (2010). Les densités les plus importantes se trouvent dans la capitale Tiranë, mais également dans les trois grandes villes Elbasan[1], Korçë et Fier, si l'on en croit les chiffres de l'association Amaro Drom. Selon Koinova (2000 : 7), il faut également y ajouter les villes de Durrës et Berat. Selon Hackaj (2014), 1005 Roms ont été recensés officiellement en 2011 dans le district de Korçë, mais la réalité approcherait plutôt les 800 familles, soit 23 000 personnes. Ce nombre inclut cependant les Égyptiens (cf. section 1.4, p. 26), et les villes de Pogradec, Pojan, Maliq et Bilisht. Dans les districts de Berat et de Vlorë se sont recensées respectivement 202 et 282 personnes en tant que Roms – mais il y aurait en réalité respectivement 2 400 et 500 personnes selon l'auteur.[2]

Les Roms vivent soit dans des quartiers urbains spécifiques, soit en périphérie des villes, soit dans des villages proches des grandes villes. Si certains vivent dans des maisons ou des appartements individuels, beaucoup vivent encore dans des conditions très précaires rappelant la description de Hasluck (1938a : 52-54), manquant d'eau et d'électricité, en périphérie des villes et sans accès à un moyen de transport. Le changement de régime s'est accompagné de nombreuses expulsions des logements qu'habitaient les Roms et qui furent « restitués » aux Albanais. Certaines familles allaient d'expulsion en expulsion (Koinova 2000 : 16). Lors de mon séjour, j'ai vu des familles vivre dans des baraquements qui rappellent les bidonvilles français.[3] C'est l'un des nombreux effets de la chute du régime hoxhiste et de la pauvreté accrue qui s'est abattue sur les Roms.[4] La chute du régime a été catastrophique pour le niveau de vie, le taux d'activité, le niveau de salaire, la santé, le logement[5], etc. Beaucoup de Roms n'ont pas accès aux services sociaux de base dans le domaine de la santé et des allocations chômage, parfois parce que cet accès leur est refusé, ouvertement ou non, par l'administration. Une autre raison est qu'un grand nombre d'enfants ne sont pas inscrits à l'état civil (Gëdeshi et Miluka 2012 : 22) depuis que l'inscription est devenue payante et hors de portée de la population la plus pauvre. Cela prive en principe ces enfants d'un accès au système de santé, d'éducation, et aux aides sociales. Une stratégie a consisté, pour beaucoup de mères enceintes qui en avaient la possibilité, à émigrer en Grèce ou en Italie le temps de l'accouchement, afin d'offrir à leurs enfants une existence légale et un accès aux droits sociaux albanais, ainsi que des papiers d'un pays de l'Union Européenne. Selon European Roma Rights Center (2010) enfin, seulement 76 % des Roms sont vaccinés contre les principales maladies infectieuses (polio, diphtérie, tétanos et coqueluche), contre 93 % des Non-Roms. Mecaj (2016) dresse un tableau des programmes d'aide humanitaire des

[1] Confirmé par Sula Raxhimi (2015 : 39).

[2] Ces chiffres me paraissent largement sous-estimés.

[3] « Au fil du temps, on constate que le nombre de sites illicites à l'allure de bidonvilles qui disparaissent et réapparaissent ailleurs s'accroît et se propage autour de la périphérie de capitale – telles que ceux de Shkoza, du lac, de la gare, d'Artistike, de Selita, Sharrë, etc. –, abritant chacun entre 100 et 300 individus. » (Sula Raxhimi 2015 : 235)

[4] Sula Raxhimi (2015 : section 6.1) décrit la vie dans un campement de Tiranë. « Tout semble très loin d'ici, la ville, les institutions, la vie décente. Ce lieu et ses habitants semblent être abandonnés dans leur lutte contre la misère et l'injustice. De temps à autre, des journalistes et des ONG leur rendent visite. Mais une fois leurs reportages publiés et leurs images choquantes transmises à la télé, tout redevient comme avant ; la vie dans le camp reprend le dessus avec tous les défis et les difficultés du quotidien. » (Sula Raxhimi 2015 : 194). Les bidonvilles se sont multipliés autour de la capitale, cœur économique du pays, en raison de la libéralisation du pays, provoquant la fermeture ou la privatisation des entreprises et le chômage (Sula Raxhimi 2015 : 235).

[5] Voir à ce sujet Andoni et Orgocka (2013) et Sula Raxhimi (2015 : 235, 237-239, 243-245).

ONG en Albanie dans les années 2000 et 2010, qui n'ont malheureusement pas apporté le progrès souhaité dans les domaines de l'éducation, de l'emploi, du travail et du logement. Elle évoque la carence de données fiables et le manque de compréhension des traditions et de l'histoire roms pour parvenir aux progrès en question. On peut cependant dénoncer la compromission de l'ensemble des organisations « non-gouvernementales »[1].

1.2. Éléments sociologiques

1.2.1. Instruction

Si l'instruction est théoriquement obligatoire pendant huit ans (cycle primaire inférieur et supérieur, cf. Tableau 2), de 6 à 14 ans, peu de Roms achèvent réellement ce cycle.

Tableau 2 : Organigramme du système scolaire albanais

Cycle	Nom	Durée	Age	Statut
pré-primaire	*parashkolla* (« pré-école »)	3 ans	3 – 6 ans	facultatif
primaire	*shkollë tetëvjeçare – cikël i ulët* (« école de 8 ans – cycle inférieur »)	4 ans[2]	6 – 10 ans	obligatoire
	shkollë tetëvjeçare – cikël i lartë (« école de 8 ans – cycle supérieur »)	4 ans	10 – 14 ans	
Passage du *provime lirimi* (« examen de fin d'études »), en vue du *dëftesë lirimi* (« certificat de libération »)				
secondaire	*shkollë e mesme e përgjithshme* (« école moyenne générale »), communément appelée *gjimnaz* (« lycée »)[3]	3, 4 ou 5 ans	14 – 18 ans	facultatif
	shkollë e mesme profesionale (« école moyenne professionnelle »)		14 – 17 ans	
	shkollë teknike (« école technique »)		14 – 19 ans	
Passage du baccalauréat, *dëftesë pjekurie* (« certificat de maturité »), communément appelé *matura*				
tertiaire	*universitet* (« université ») *institute i lartë* (« institut supérieur ») *akademi* (« académie »)	2 à 6 ans	± 18 ans – ± 24 ans	facultatif
Passage d'un diplôme, *diplomë* : *dëshmi specializimi* (« certificat de spécialisation ») ou *doktor* (« docteur »)				

[1] « Les ONG locales […] sont pour la plupart créées et quasi entièrement financées par les donateurs et les gouvernements étrangers. Cette dépendance vitale expliquerait d'ailleurs en partie la révérence envers les agences étrangères diplomatiques et de l'aide au développement, ainsi que l'absence de toute critique quant à leur rôle dans le pays. Les ONG roms ou celles qui travaillent pour "résoudre les problèmes" des Roms, ne font pas exception à la règle, elles sont toutes, à ce jour, soutenues financièrement par des bailleurs de fonds, des agences et des gouvernements étrangers, à quelques exceptions près. » (Sula Raxhimi 2015 : 248)

[2] Depuis la réforme de l'éducation de 2004, le cycle inférieur, équivalent à l'école primaire, dure cinq ans au lieu de quatre, ce qui porte à neuf le nombre d'années de scolarité obligatoire, au lieu de huit auparavant.

[3] Depuis la réforme de l'éducation de 2004, le lycée dure trois ans au lieu de quatre.

Beaucoup d'enfants ne suivent que la première moitié du premier cycle obligatoire, de 6 à 10 ans, voire une partie seulement ou rien du tout, et ne savent donc pas ou peu lire et écrire. L'État ne les encourage pas puisqu'il n'existe pas de subvention ou de bourse spécifique pour aider les Roms à suivre l'école, le lycée ou l'université. Selon Koinova (2000 : 29-30), 40 % des enfants roms suivent la première partie du cycle primaire et 3 % la deuxième partie du cycle primaire (obligatoire) – alors que les statistiques officielles parlent de 100 % d'assiduité scolaire pendant le cycle primaire obligatoire. Moins de 1 % des jeunes Roms fréquentent le lycée, l'université ou des écoles supérieures, alors que les Albanais les fréquentent à 72 %. Les jeunes Roms qui font des études supérieures deviennent souvent travailleurs sociaux ou éducateurs (Koinova 2000 : 30). Ceux qui réussissent à étudier ou à travailler à l'étranger ne reviennent alors plus vivre en Albanie. L'European Roma Rights Center (2010) donne des chiffres plus encourageants : environ 52 % des Roms n'ont pas du tout été à l'école, 18 % ont fait quelques années d'école obligatoire et 14 % ont achevé le cycle primaire obligatoire (huit années de scolarité). 3% ont terminé le lycée et 4% ont obtenu un diplôme d'études supérieures. Les femmes sont moins instruites que les hommes : 56 % n'ont pas été à l'école, et 11% n'ont suivi que quelques années d'école. Sula Raxhimi (2015 : 234) dit que l'analphabétisme « s'est notamment accentué ces deux dernières décennies, touchant environ 40 % des Roms, par rapport à seulement 1,6 % de la population majoritaire. En 1981, 639 enfants roms de 6 à 16 ans – sur un total de 2 363 enfants roms –, n'étaient pas scolarisés, soit 27%, alors qu'en 1982, environ 853 individus – tous âges confondus – restaient en dehors du système d'éducation, n'allaient pas à l'école ou l'avaient abandonnée. » Deux raisons majeures sont invoquées pour expliquer l'abandon de l'école (Koinova 2000 : 16). D'une part, les enfants souffrent de discrimination spécifique et de mauvais traitements (coups, insultes) de la part des élèves non-roms ainsi que des enseignants (raison déjà invoquée par Hasluck 1938b). D'autre part, l'école est souvent à une grande distance du domicile, en périphérie des villes ou en zone rurale. Aller à l'école, presque toujours à pied, représente donc de la fatigue et un danger – des cas de kidnapping ont été avérés dans les années 1990. Or l'absence de scolarisation et la scolarisation non adéquate, qui empêchent l'accès au marché du travail et aux opportunités, sont l'une des causes de la pauvreté (Gëdeshi et Miluka 2012, Shanaj, Hallkaj et Cuninghame 2015). Or plus de 78% des Roms se déclarent « très pauvres » en 2011 (Sula Raxhimi 2015 : 234-235).

Ces deux facteurs existaient déjà dans l'entre-deux-guerres et sous le régime hoxhiste : la chute de celui-ci a en fait aggravé la situation. Les Roms furent touchés de plein fouet par le chômage, le délabrement du système social, notamment du système de prestations sociales et de santé. La moyenne d'illettrisme est alors passée de 39,6 % à 62,2 % chez l'ensemble des Roms, si l'on compare la génération des 21-40 ans à celle des 7-20 ans (Tableau 5). Les jeunes Roms d'aujourd'hui font en moyenne deux années d'école de moins que leurs parents (4,02 années au lieu de 6,22 années). L'absentéisme scolaire a fortement augmenté durant les années 1990 et 2000 (Koinova 2000 : 12) en raison du chômage des parents, qui ne pouvaient pas assumer le coût de la scolarité de leurs enfants selon (Koinova 2000 : 6). Les parents déclarent en effet ne pas pouvoir envoyer leurs enfants à l'école parce qu'ils ne peuvent pas leur acheter des habits décents, les livres et le matériel scolaire requis, ou encore parce qu'ils ont besoin des enfants pour les aider à s'occuper des plus petits, à prendre en charge le travail domestique, ou encore pour les aider à subvenir aux besoins de la famille (cf. les réponses au questionnaire présentées par De Soto, Beddies et Gedeshi (2005 : 191). En outre, le système scolaire ne prévoit pas de cours de langue romani ni en langue romani, qui est la langue

maternelle et d'usage quotidien de la minorité – alors qu'il prévoit des cours en grec dans une vingtaine d'écoles pour la minorité grecque selon Koinova (2000 : 26). On peut invoquer un cinquième argument, d'ordre culturel : l'âge du mariage plus précoce chez les Roms que chez la majorité non-rom (Koinova 2000 : 28), pour les garçons comme pour les filles. Beaucoup de jeunes sont amenés à fonder un foyer entre 14 et 19 ans, soit pendant le cycle secondaire, non obligatoire. Les filles qui vont à l'école peuvent notamment être retirées de l'école dès le début de la puberté, vers 12 ans, soit au début de la deuxième partie du cycle primaire obligatoire.[1]

1.2.2. Activité professionnelle

D'après ce que j'ai pu observer, il n'y a pas ou plus de profession associée en particulier aux différents groupes aujourd'hui en Albanie. La chute du régime hoxhiste à la fin des années 1990, les privatisations qui suivirent, l'affaire dite des Pyramides et la crise financière de 1997 qui ont provoqué la guerre civile, furent particulièrement désastreuses pour les Roms et pour les Non-Roms défavorisés. Ils furent les premiers à être licenciés des entreprises industrielles et agricoles et des services d'État. Plus que les autres ils affrontèrent le chômage de masse, et notamment les jeunes. L'absence ou le manque d'instruction scolaire et de qualification professionnelle, ainsi que la discrimination à l'embauche (ressentie par 30 % des Roms selon De Soto, Beddies et Gedeshi 2005 : 67-68), produisent un taux de chômage beaucoup plus élevé que dans la population majoritaire : 80 à 90 % chez les Roms contre une moyenne officielle nationale de 18 % en 1996 selon Koinova (2000 : 16), 71 % chez les Roms contre une moyenne officielle nationale de 16 % fin 2002 (De Soto, Beddies et Gedeshi 2005 : 66).[2] Selon ces derniers, 93% des Roms actifs l'étaient sans contrat de travail (De Soto, Beddies et Gedeshi 2005 : 66). Selon European Roma Rights Center (2010), le revenu moyen d'un foyer rom est de 68 euros mensuels, tandis que le revenu moyen d'un foyer non-rom du même quartier serait de 174,50 euros.

Les emplois, souvent temporaires, qu'ils occupent, sont principalement dans les secteurs de l'agriculture et du bâtiment, et non plus de l'artisanat comme dans les années 1930. L'embauche se fait souvent au noir et à la journée, par exemple chez les porteurs et dockers (De Soto, Beddies et Gedeshi 2005 : 70). Certaines activités traditionnelles ont cessé aujourd'hui d'être exercées en raison de la modernisation de la société : vannerie, chaudronnerie, maquignonnage. D'autres activités, qui n'étaient plus d'actualité sous le régime, voire prohibées en raison de l'interdiction de sortie du territoire, sont à nouveau pratiquées, telles que l'emploi saisonnier dans l'agriculture, en Albanie ou en Grèce. De Soto, Beddies et Gedeshi (2005 : 13) parlent d'une « ré-émergence de beaucoup d'aspects de [l']identité culturelle » des Roms et des Égyptiens, provoquée par la chute du régime. D'autres enfin, qui se pratiquaient sous le régime, se poursuivent, ainsi la musique ou le commerce de proximité, notamment la vente de vêtements d'occasion sur les marchés. Ces vêtements usagés

[1] Pour une approche anthropologique de la minorité rom d'Albanie, on pourra consulter le travail de Sula Raxhimi (2015). Sa section 2.3.1 propose une synthèse de l'anthropologie européenne des Roms, notamment des Balkans.

[2] Sula Raxhimi (2015 : 234) écrit que « le taux de chômage pour les Roms s'élève à plus de 50% – et dans certaines poches urbaines à plus de 73% –, alors que ce même taux était de seulement 11% en 1976 et de 29% en 1981. Actuellement, plus que 92% des Roms travaillent au noir dans des secteurs informels de l'économie. »

viennent d'Albanie, de Macédoine, mais aussi de Turquie, de Grèce et d'Italie. Les revenus tirés de cette activité ont cependant nettement baissé à la chute du régime, en raison de la diminution de la demande (due à la hausse de la pauvreté) et de l'augmentation de la concurrence par d'autres pratiquant la même activité (selon De Soto, Beddies et Gedeshi 2005 : 68-69, 72). La mendicité est pratiquée à l'occasion, pour de l'argent, des habits et de la nourriture, surtout par les femmes et les enfants, de même que la collecte d'objets en métal pour la revente au poids (De Soto, Beddies et Gedeshi 2005 : 71). Les femmes gagneraient également de l'argent en disant la bonne aventure en ville, ou encore en vendant leur sang à l'hôpital (De Soto, Beddies et Gedeshi 2005 : 71-72). Les Roms occupant des professions dites intellectuelles ou supérieures, telles qu'enseignant, médecin ou juriste sont donc des exceptions notables et marquantes, qui font la fierté de leur famille et de la communauté. Le Tableau 3 résume la situation socio-économique des Roms actuellement en Albanie.

Tableau 3 : Quelques données sociologiques comparatives (Sula Raxhimi 2015 : 236)

Description	Les Roms	La population albanaise
Analphabétisme	40,3 %	1,6 %
Nombre moyen d'années de scolarisation	5,6	> 9
Fréquentations des 9 premières années d'école	16,1%	57 %
Taux d'inscription à l'école	44,1%	> 90 %
École secondaire	2,1%	21,7 %
Taux de chômage	> 50 %	13,5 %
Travail indépendant	> 60 %	N/A
Pauvreté	> 78 %	12,4 % (26 % régions reculées)

Source : INSTAT – Recensement de la population albanaise 2011; Études de l'UNICEF (2011) et de PNUD (2012)

1.2.3. Discrimination

Le travail de Sula Raxhimi (2015) dépeint la dynamique d'exclusion, de violence et de discrimination raciale à l'œuvre de nos jours en Albanie, et héritée depuis l'Empire ottoman. Les différentes formes de violence dont souffrent les Roms ont déterminé leur statut et leur place dans la société albanaise depuis des générations, sur les plans social, économique et politique – ce qui « affecte leur éducation, leur mode de vie et leur statut en retour » (Sula Raxhimi 2015 : 260). Nous avons vu que les Roms se disaient discriminés ou maltraités à l'école, par les élèves et/ou par les enseignants, ce qui serait l'une des causes de l'absentéisme scolaire. Quelque 30 % des Roms ressentent le racisme comme la raison principale du chômage (De Soto, Beddies et Gedeshi 2005 : 67-68). Le chômage, la précarité ainsi que l'exclusion sociale, éducationnelle et sanitaire touchent tout particulièrement les Roms depuis la chute du régime, ce qui explique la nostalgie qu'éprouvent aujourd'hui beaucoup d'adultes. Les Roms font en revanche l'objet d'un grand nombre de stéréotypes dégradants, étant entre autres considérés comme « pauvres, sales, stupides, bruyants et voleurs », « moralement inférieurs » voire violents (Koinova 2000 : 17-18). Ces images négatives sont largement répandues par les films, téléfilms et articles de journaux. La pauvreté qui les frappe les marginalise et renforce les stéréotypes, notamment lorsque certains tombent dans la délinquance (Courthiade 1995 : 14). La préservation de

la culture traditionnelle sert parfois de prétexte aux autorités pour leur refuser l'accès aux services sociaux et sanitaires.

Selon (Koinova 2000 : 14), les Roms sont également victimes de discrimination de la part de la police, les autorités municipales ainsi par l'armée au cours du service militaire. La première s'est montrée de nombreuses fois coupable de violence et harcèlement envers les Roms dans les années 1990 : passages à tabac en public, extorsion de fonds[1], arrestation et/ou détention arbitraire contre rançon, mauvais traitements. Les secondes refuseraient ou feraient du chantage à l'attribution des prestations sociales, médicales et sanitaires, comme l'alimentation en eau ou le raccordement à l'électricité. Pour avoir accès à des soins, les Roms doivent souvent verser d'importants pots-de-vin, peut-être plus importants que ceux des Non-Roms. Enfin, les Roms seraient victimes de mauvais traitements et considérés de manière inégalitaire dans l'armée.

Les méfaits de la police et les actes criminels isolés de la part d'individus seraient cependant moins nombreux que dans la plupart des autres pays d'Europe de l'Est où les agressions sont plus fréquentes, plus sévères et plus ostensiblement racistes. Les études s'accordent à dire que les Roms albanais ne sont pas victimes de violences raciales ouvertes – à moins que celles-ci ne soient cachées et ignorées de l'opinion publique et des organisations internationales – mais plutôt de discrimination larvée. En outre, les inégalités sociales sont moindres en Albanie qu'ailleurs, c'est-à-dire que le différentiel entre les problèmes socio-économiques (pauvreté, chômage, logement insalubre, manque d'accès aux soins, malnutrition…) que rencontrent les Roms et la moyenne de la population non-rom est moins important que dans d'autres pays. L'Albanie est le deuxième pays le plus pauvre d'Europe[2]. Même si les Roms y souffrent d'extrême pauvreté, de marginalisation et de discrimination, d'où le recours à la migration économique hors du pays, cela fait paradoxalement de l'Albanie un pays considéré comme relativement « bon à vivre » pour les Roms (selon Koinova 2000 : 19-31).

Pour plus d'informations sur la situation sociale et économique des Roms albanais, on pourra consulter l'étude menée pour la Banque Mondiale par De Soto, Beddies et Gedeshi (2005), celle menée par le Programme des Nations Unies pour le développement par Gëdeshi et Miluka (2012), la synthèse de UNDP (2012), celle menée par Dobrushi 2013) pour l'Open Society, celle de Simon, Galanxhi et Dhono (2015) pour la Direction suisse du développement et de la coopération, celle de Friedman (2015) pour le Programme des Nations Unies pour le développement. La synthèse de Hackaj (2014) concerne plus spécifiquement les conditions de travail et de chômage des Roms de Korçë, Berat et Vlorë. La multiplication des états des lieux, enquêtes et compte-rendu montrent un réel intérêt des autorités internationales. « L'État n'est en effet pas le seul acteur sur la scène politique en Albanie, du moins pas en ce qui concerne la politique envers les Roms. Les acteurs les plus puissants et présents sont les organisations politiques, onusiennes, les diplomates et les ONG (Sula

[1] L'extorsion de fonds peut s'expliquer par la croyance populaire selon laquelle les Roms sont riches, notamment ceux travaillant à l'étranger comme saisonniers (Grèce, Italie) donc détenant des devises fortes.

[2] En 2014, le PNUD a classé l'Albanie 85e mondiale pour son IDH, ex-aequo avec la Bosnie-Herzégovine et l'Arménie, et derrière la Macédoine 81e, la Turquie 72e, la Serbie 66e, la Bulgarie 59e, la Roumanie 52e, le Monténégro 49e, la Croatie 47e et la Grèce 29e. Seule la République de Moldavie est plus mal placée que l'Albanie : 107e (Jahan et Jespersen 2015 : 220-223).

Raxhimi 2015 : 246), ainsi que les ONG locales et les associations roms elles-mêmes.[1] En réalité, on peut considérer que la multiplication des organisations montre que les Roms sont considérés « comme étant en dehors de la société, comme un problème sociétal, comme une communauté à intégrer » (Sula Raxhimi 2015 : 256), comme s'ils ne faisaient pas partie de la société albanaise.

1.2.4. Korçë

Il n'existe pas d'étude spécifique aux Roms de Korçë. Il est néanmoins possible de réunir quelques informations historiques et sociologiques. Le militaire géologue français Jacques Bourcart dresse une courte description des Roms de cette région au début du XXème siècle.

> « Comme dans presque tous les Balkans, les Confins Albanais avaient une population tsigane soit nomade, soit demi-fixée. Forgerons, ouvriers agricoles, moissonneurs, petits artisans, ils habitent sous la tente ou plus généralement dans de véritables gourbis qui, soit forment des quartiers spéciaux (« Katavarosh » à Korça), soit sont groupés autour des maisons des beys, dont les Tsiganes forment une véritable « clientèle ». Les Tsiganes musulmans ou orthodoxes officiellement, sont en réalité très indifférents au point de vue religieux ; les femmes se rendent indistinctement, une fois l'an, au pélérinage de Santa Marina sur le rocher de Llënga ; tous vénèrent le monastère de Saint-Naoum sur le lac d'Ohrida.
>
> Fort polyglottes, ils parlent rarement leur langue propre devant les étrangers. Ce sont des travailleurs solides, résistants et sobres, indispensables à la vie du pays, utilisables surtout dans de mauvaises conditions climatériques (Mysakja, lac Maliq). Musiciens de toutes les fêtes, il n'est pas de cérémonie sans eux en Albanie.
>
> Les Tsiganes sont groupés dans leurs quartiers spéciaux, autour de chefs, que l'administration française a pris comme intermédiaires pour agir sur eux. Malgré les traces de sang étranger dues au « droit du seigneur », ils gardent une physionomie anthropologique très particulière. » (Bourcart 1922 : 276)

En 1870, Korçë comptait près de 10 000 habitants, au début du XIXème siècle environ 15 000, et lors du recensement de 1923 quelque 25 598 habitants. Aujourd'hui, les trois principales villes peuplées de Roms seraient, après la capitale Tiranë, Elbasan, Korçë et Fier (selon les chiffres de l'association Amaro Drom et European Roma Rights Center (2010).

Katavarosh désigne aujourd'hui un quartier de Korçë. Signifiant « en deçà de la ville », il provient de *varosh*, ville et *kata*, sous. Notons que *varosh* est emprunté au hongrois *város*, cité fortifiée, terme passé dans les dialectes roumains et sud slave (il existe plusieurs communes Varoš en ex-Yougoslavie). *Varosh* était autrefois le nom d'un certain quartier, le vieux centre, les premières maisons à avoir été construites au pied de la forteresse. Korçë comprenait autrefois un château, mentionné par les sources ottomanes en 1431. On ignore toutefois son emplacement exact et il n'en reste rien[2], si ce n'est la preuve indirecte de l'étymologie. Il s'agissait d'un bâtiment à usage

[1] « Elles représentent un nouveau secteur de l'économie. Ainsi, entre 2000 et 2008, environ 800 ONG étaient enregistrées légalement en Albanie, alors que seulement la moitié […] menait des activités. Aujourd'hui, on compte 8 449 ONG enregistrées en Albanie. » (Sula Raxhimi 2015 : 248)

[2] Des morceaux de remparts ont cependant été mis à jour lors d'excavations en 1923, 1937 et 2000.

strictement militaire ; la ville proprement dite était le *varosh*, quartier situé sous et autour du fort. Lorsque le *varosh* devint trop peuplé, la ville s'étendit en périphérie. Celle-ci se nomme alors *katavarosh*, *varosh e poshtme*, « quartier sous le varosh », ou encore *varosh i ri*, « nouveau varosh ».[1] Le Katavarosh de Korçë constituait la périphérie immédiate du vieux centre, autour de l'église Sainte Marie. C'est là que vivaient les Roms de confession orthodoxe à la fin du XIXème et au début du XXème siècle. Ceux de confession musulmane habitaient un quartier en marge de la périphérie, nommé Radanec. Durant la Première Guerre Mondiale, Korçë fut occupée par l'Entente : l'armée et l'administration grecque de 1912 à 1916 (qui rattache la ville à sa région d'Epire du nord), puis l'armée et l'administration française de 1916 à 1920. Sous l'occupation grecque, les Roms de Katavarosh furent expulsés de leurs maisons et durent créer un nouveau quartier, *Kulla e Hirit*, « la Tour de la Grâce » – expression homonyme, en albanais, de « la Tour des Cendres ».

On sait depuis Hasluck (1938b) que, si les Roms d'Albanie sont pour la plupart musulmans, les régions de Korçë et de Gjirokastër font exception car elles comptent une majorité d'orthodoxes – pratiquant les mêmes rites dans les mêmes églises que les Non-Roms.

Du point de vue professionnel, les Roms de Korçë sont semblables à ceux des autres régions d'Albanie. Mann (1933a : 1-2) rencontre à Tiranë deux groupes de Roms parlant deux variétés différentes : celui du nord (Shkodër) et celui du sud (Korçë). Ceux du nord ont migré selon lui de Serbie et vivent de la chasse aux chiens sauvages et du maquignonnage. Ceux du sud ont migré de Grèce et de Macédoine – vraisemblablement celle grecque, Florina et Salonique ; ils vivent de maquignonnage, d'artisanat (vannerie, objets décoratifs) et de mendicité. De nombreuses femmes sont domestiques[2] dans les riches demeures de Non-Roms. Les Roms du sud rencontrés par Mann appartiennent vraisemblablement à deux groupes différents, Arlis et Čergars II (cf. section 1.3, p. 16). Ils viennent de Korçë, Pogradec, Florina et Salonique. Il les décrit comme « les plus industrieux » : les hommes vendaient des chevaux et des mules, confectionnaient des paniers et des fleurs en papier (Mann 1933a : 2), tandis que les femmes et les enfants mendiaient (Mann 1935 : 174). Il rencontre un couple ayant voyagé à travers toute l'Europe et même aux États-Unis où l'homme, Neğipi, avait travaillé comme clarinettiste. Selon Margaret Hasluck (note de (Mann 1935 : 175), Neğipi fait probablement partie de la population musulmane de la Macédoine grecque, expulsée en 1924 suite à l'« échange de population » arrangé entre les États grec et turc. Seuls les musulmans de « nationalité » albanaise avaient pu demeurer, mais non les musulmans roms et turcs. Certaines familles choisirent de migrer vers la Thrace orientale (turque), tandis que d'autres choisirent de rester dans la même région, à quelques kilomètres, en passant la frontière albanaise. En Albanie vivaient en effet beaucoup de Roms, de musulmans et de Roms musulmans.

Voici enfin la situation des Roms dans le district de Korçë au XXIème siècle, telle que la décrivent (De Soto, Beddies et Gedeshi 2005 : 231-232) :

[1] Ailleurs et en l'absence de forteresse, on parle de *mahalla*, quartier. La périphérie nouvelle se nomme alors *yeni mahalla*, nouveau quartier.

[2] Vaux de Foletier (1981) l'atteste : « J'ouvre une parenthèse, ayant visité l'Albanie et ayant été reçu par les Courville [à Korçë] en 1936. Au temps du roi Zog, la coutume ne permettait pas à une femme proprement albanaise de servir dans une maison où il y avait un homme autre que son père, son frère ou son mari ; les femmes tsiganes acceptaient volontiers cet emploi ».

« *The district is relatively prosperous, with high education levels. Most economic activity is centered in the city of Korça, its administrative center. The city (pop. 55,130; INSTAT 2001) is home to cultural and historical monuments. The district's major economic resources are agriculture, scattered small businesses, and commerce. The district's unemployment rate is 22 percent (INSTAT 2002). Remittances comprise a large portion of family incomes.*

Most Roma in Korça are Karbuxhinj who refer to themselves as "Erlinj."[1] *They migrated from Turkey to Folorina* [sic], *Greece, before moving to Korça in the beginning of the 20th century. In the city, Roma families live in dilapidated and overcrowded homes concentrated in the "Kulla e Hirit" neighborhood. Major sources of their income include used clothes sales, casual work, and state economic assistance.*

Other Roma families live in different villages throughout the district, such as Maliq, Libonik, Pojan, and Sovjan. Income sources include horse-dealing, agricultural work, used- clothes sales, casual work, and seasonal employment in villages over the Greek border. Some families recently sold land awarded to them in 1992 through state land distribution. In Maliq, unemployment is high because of the bankruptcy of the town's sole major enterprise, a sugar mill (De Soto and others 2001).

Approximately 100 Roma families live in the commune of Pojan. These Roma sold land they had acquired after 1990. According to land registry records, 24 out of 100 Roma families sold 44,000m² of land during the transition period. Families that have retained their land rely heavily on subsistence agriculture. Roma identified running water as a major problem in Pojan. According to Pojan's kryeplak, residents have experienced several months without running water. When water service is disrupted, residents travel to surrounding villages to collect drinking water. In spite of all, primary school attendance is high compared to attendance levels in other Korça district localities.

Egyptians represent approximately 20 percent of the city's population and live in two neighborhoods, respectively of Muslim and Orthodox Egyptians. Most live in dilapidated and overcrowded houses, while others camp in temporary housing provided by the local municipality. This community is also characterized by high unemployment, low school attendance rates, and prostitution. » (De Soto, Beddies et Gedeshi 2005 : 231-232)

1.3. Les groupes

Comme dans les autres pays, la société rom est structurée en groupes (Piasere 2011). Le terme utilisé par les Roms pour les désigner est emprunté à l'albanais : *fis(i)*, pluriel *fisja*. Chaque *fis* serait à l'origine une grande famille descendant d'un ancêtre mâle commun. Ce terme est aujourd'hui ambigu : il désigne aussi bien le groupe, à l'échelle du pays, qu'une famille précise (romani *familja*), dont le souvenir ne remonte pas à plus de cinq générations. De Soto, Beddies et Gedeshi (2005 : 21, 22, 222) illustrent la notion de *fis*, au sens de « famille », grâce à l'arbre généalogique de la famille qui m'a accueillie à Korçë (cf. la reproduction de l'arbre généalogique en Figure 89 et les récits en Figure 90). Elle fait partie du *fis*, au sens de « groupe », des Arlis. S'il s'agit à l'origine du même concept anthropologique, le terme renvoie aujourd'hui à deux réalités différentes : on distinguera en français la *famille* (plus ou moins élargie) et le *groupe*, se distinguant par leur variété dialectale, par leur métier, leur mode de vie, éventuellement par leur origine nationale s'il y a eu migration. Pour un

[1] « *"The word "Erli" is a Turkish word that can mean "native inhabitant" or "non-nomadic".* »

aperçu des noms des différents groupes en Europe, on pourra consulter Matras (2004 : 66-68) ou encore Courthiade (2009a).

Il n'existe pas d'étude traitant spécifiquement des groupes de Roms albanais. Les informations qui suivent proviennent d'une part des présentations généralistes de Koinova (2000) et de De Soto, Beddies et Gedeshi (2005), d'autre part de mon travail de terrain : observations et témoignages des Roms sur place (avec entretiens et récits de vie). Les groupes parlant des variétés différentes seraient au nombre de six : les Mečkars, les Arlis, deux groupes de Čergars (que l'on nommera I et II), les Bamils et les Kurtofis. Les deux premiers sont les plus importants numériquement, car ils représentent chacun au moins un tiers de la population rom albanaise. Ces différents groupes seraient répartis partout dans le pays. Comme en Macédoine pour les Kovač, les Džambaz et les Arlis, on peut dire que les trois principaux groupes albanais (Mečkars, Arlis, Čergars I) sont homogènes et distincts historiquement et socialement. Notre hôte a mentionné également un autre groupe de Roms, qui vivrait dans le sud de l'Albanie, les Konop(l)jaris. Il pourrait s'agir du même groupe que les Konopars macédoniens dont Friedman (2017) décrit la variété et que mentionne Heinschink (1978). Le terme signifie « fabriquant de corde » (du macédonien *konop*, corde). N'étant pas sûre que ce groupe réside bel et bien en Albanie, j'ai choisi de ne pas l'inclure dans la liste suivante.

D'après ce que j'ai pu observer sur le terrain, l'endogamie est souvent de mise entre Roms, mais les mariages entre personnes de différents groupes (et villes) sont normaux et fréquents. Cela assure une certaine mixité tant génétique que linguistique : beaucoup de locuteurs parlent ou ont ainsi parlé deux variétés de romani en famille durant leur enfance. Il n'existe pas non plus de religion associée à tel ou tel groupe de Roms. Les Roms albanais sont pour la plupart musulmans, à l'image du reste de la population du pays. La conversion à une autre religion, comme le christianisme orthodoxe, est un moyen d'ascension sociale sur le plan local, ou encore un moyen de faciliter la migration vers un pays chrétien (Grèce, Italie…).

1.3.1. Ethnonymes

Il convient de distinguer le nom que se donne un groupe (*endonyme*) du nom que lui donnent les autres groupes (*exonyme*). Le terme principal est *Rom*, endonyme parfois utilisé comme exonyme. *Magjup* est un terme-chapeau utilisé surtout au nord de l'Albanie et au Kosovo. Le terme *Cigan* n'existe pas dans la langue populaire et reste, selon Marcel Courthiade, confiné aux études scientifiques, à la littérature étrangère et à la musique rom hongroise. Enfin, le terme Arixhi était employé comme exonyme dans les textes et documents d'archives albanaises (Sula Raxhimi 2015 : 34).

Il existe des noms d'utilisation régionale (Koinova 2000 : 16-17) pour désigner les Roms : *Gabel·e*, étranger, utilisé en Albanie centrale (Courthiade 1988b), *Arixhi/Arixheshka*, oursier/oursière, désignant plutôt les Roms nomades selon (Courthiade 1985) utilisé au sud du pays, *Katal/Katalka* ou *Gatal/Gatalka*, diseur/diseuse de bonne aventure, utilisés respectivement au nord du pays et à Dibra (aujourd'hui en Macédoine du Nord) selon Margaret Hasluck, *Kurbat/Kurbatka*, migrant/migrante,

étranger/étrangère, exil, ou *Skinit·e*, habitant de tente[1], (Hasluck 1938a : 50), utilisé à Korçë pour désigner les Čergars II, *Qifto* (« Égyptien ») utilisé à Gjirokastër. Hasluck (1938a : 50-51) donne d'autres noms, dont il est difficile de savoir s'ils s'appliquent à tous les Roms ou à un seul groupe : *Çerget(ar)·e*[2], habitant de tente, à Berat, *Muhaxhir·e*, émigrant, exilé, réfugié, à Libohovo, *Shoshatar·e*, fabriquant de tamis, à Kolonja.

La distinction endonyme/exonyme s'applique également aux groupes de Roms. Le Tableau 4, élaboré et discuté avec Marcel Courthiade, permet de visualiser les noms des différents groupes. Il apporte une distinction supplémentaire entre endonyme *primaire*, nom donné par les membres d'un groupe « pour se désigner en utilisant un terme provenant de leurs propres ressources lexicales » (en **bleu** dans le tableau), et endonyme *secondaire*, nom donné par les membres d'un groupe « pour se désigner en utilisant un terme provenant d'un lexique étranger » (en **vert**). En noir apparaît l'exonyme, nom donné par les étrangers au groupe pour désigner les membres dudit groupe.

Ce tableau présente cinq groupes : les **Mečkars**, les **Arlis** (nommés ici *Kabo-Erli*), les Čergars de Shkodra (nommés dans le tableau *Škodràni*[3], mais que l'on préférera nommer **Čergars I**), les Čergars de Peqin et Permet (également nommés *Rupane*, que l'on nommera **Čergars II**), les Bamils (nommés ici **Bamìþi**). Marcel Courthiade ne prend pas en compte l'existence des **Kurtofis**. Il mentionne en outre un prétendu groupe nommé **Vakërde**, « d'une rare violence et très primitifs, sédentaires, mais il semble que ce soit plutôt une légende urbaine qu'une réalité ethnologique – une sorte de repoussoir identitaire ».

[11] Terme grec utilisé en Grèce pour désigner les « *Arvanitovlachoi* », nommés encore Čobani ou Vlachs : ce sont des pasteurs et commerçants nomades parcourant plusieurs régions des Balkans et parlant une variété romane (principalement l'aroumain). Selon Hasluck (1938a : 51), cette population pratiquait, sous l'Empire ottoman, la migration saisonnière entre les montagnes macédoniennes l'hiver, les plaines de la côte albanaise l'été. Ils se retrouvèrent « enfermés », dans les années 1930, du côté albanais ou du côté macédonien.

[2] Une déformation de *Čergar* selon Marcel Courthiade.

[3] Ces Roms étaient, selon Courthiade (1988b), initialement nombreux à Shkodër, mais vivraient quasiment tous à Tiranë aujourd'hui.

Tableau 4 : Endonymes et exonymes des Roms, Égyptiens et Albanais

<table>
<tr><th colspan="2">X nomme Y</th><th colspan="7">Y</th></tr>
<tr><td rowspan="10">X</td><td></td><td>Albanais</td><td>Égyptien</td><td>Mećkàri</td><td>Škodràni</td><td>Kabo-Erli</td><td>Rupano</td><td>Bamìþi</td></tr>
<tr><td>Albanais</td><td>Shqiptar·ë</td><td>Jevg·ë
Evgjit·ë</td><td colspan="5">Arixhi/Arixhofkë, Gabel·ë, Kurbat/Kurbatkë[1] [Korçë]</td></tr>
<tr><td>Égyptien</td><td>Gaxhoj/Gaxhie
[iron.], Ćoće[2] [arch.]</td><td>Jevg·e</td><td colspan="5">Gabel·e</td></tr>
<tr><td>Mećkàri</td><td>Gaʒo/Gaʒi, Das</td><td>Rlìa/Rìcka</td><td>Mećkàri</td><td>Ćergàri, Škodràni</td><td rowspan="2">Ka(l/r)buʒi[3]</td><td rowspan="3">Ćergàri</td><td>Bamìþ/r/li</td></tr>
<tr><td>Škodràni</td><td rowspan="2">Xoraxaj/Xoraxajni[4],
Gaʒo/Gaʒi</td><td>Arlìa</td><td rowspan="5">Mećkàri/
Mećkàrka</td><td>Rrom/Rromni, (Ćergàri, Škodràni)</td><td>Bamìþ/r/li (?)[5]</td></tr>
<tr><td>Kabo-Erli</td><td>Ćibano/
Ćibanka</td><td rowspan="3">Ćergàri/Ćergàrka</td><td>Erlìa[6]</td><td rowspan="3">Bamìþ/r/li</td></tr>
<tr><td>Rupano</td><td rowspan="2">Gaʒo/Gaʒi</td><td>Ćibano/
Ćibani</td><td>?</td><td>Rrom/Rromni (Ćergàri)</td></tr>
<tr><td>Bamìþi</td><td>Jifto[7]</td><td colspan="2">Ćergàri</td><td>Rrom/Rromni</td></tr>
</table>

[1] *Kurbat* n'est pas un terme péjoratif, au contraire d'*Arixhi* et de *Gabel*.

[2] Ce terme est en fait utilisé par les Égyptiens pour désigner tous les étrangers au groupe.

[3] Ce terme est souvent prononcé ou perçu comme péjoratif, alors qu'il signifie simplement « fabriquant de tamis » en turc. Certains Arlis ont adopté ce terme pour se désigner, soit par prestige, soit pour s'adresser à des Non-Arli (cf. l'interview dans le corpus spontané p. 576).

[4] Selon Marcel Courthiade, les Škodràni et les Kabo-Erli, c'est-à-dire les Čergars II et les Arlis, ne nommeraient pas les Albanais majoritaires *Gaʒe* lorsqu'ils parlent entre eux, mais *Xoraxane*. Le terme *Gaʒe* ne serait utilisé qu'en conversation avec les autres Roms. Notons que cette affirmation contredit Heinschink (1978 : 12) selon lequel *Khorane* est employé pour s'auto-désigner et se démarquer des Roms chrétiens, nommés *gadže*. J'ai cependant ajouté le terme *Gaʒe* dans le tableau, car je l'ai bel et bien entendu sur le terrain – mais je suis précisément non-rom. En revanche, j'ai pu entendre *K(h)orane* seulement pour désigner des Turcs ou turcophones.

[5] Selon Marcel Courthiade les Bamils sont mal connus des autres Roms : les Čergars II (Škodrànë) auraient ignoré leur existence jusqu'à il y a une dizaine d'années.

[6] J'ai pu entendre les formes *Rli, Erli, Arli*.

[7] Le terme *Jifto* serait également utilisé par les Bamils pour désigner les autres groupes de Roms albanais.

On constate que tout le monde en Albanie fait de nos jours une distinction terminologique entre Roms et Égyptiens. Les Égyptiens ont des dénominations variées : quasiment tous leur donnent un exonyme différent. Les Albanais et les Égyptiens en revanche ne font aucune différence entre les différents groupes de Roms : les noms mentionnés en section 1.3.1, p. 17, varient en fonction des régions et non des groupes désignés. Selon Marcel Courthiade, les paysans albanais feraient la différence entre « leurs » Roms locaux, dont l'endonyme spécifique est connu, et les autres.

Les noms de groupes de Roms ne sont pas hérités, mais empruntés à des langues de contact (turc, albanais), si besoin suffixés en romani, albanais, slave ou turc. Les Mečkars sont le seul groupe dont l'endonyme est identique à l'exonyme – au contraire des Arlis qui portent plusieurs exonymes différents de leur endonyme. On notera que le terme « arli », sous ses différentes variantes phonétiques, est à la fois l'endonyme du groupe Arli (*Erli*), auquel appartient la famille de mon hôte à Korçë, et l'exonyme utilisé par les Mečkars et les Škodrànă/Čergars I pour désigner les Égyptiens (*Arli*, *Rli*), équivalent donc de l'albanais Jevg·ë/Evgjit·ë.

Les Bamils, les Škodrànă/Čergars I et les Rupane/Čergars II sont appelés respectivement Bamils et Čergars par tous les Roms… sauf par eux-mêmes : leur endonyme est tout simplement *Rom*. Selon Marcel Courthiade, les Bamils vivent surtout au sud du pays et connaissent mal le reste du pays et les autres Roms : c'est pourquoi ils appellent indifféremment *Čergars* tous les autres groupes, hormis les Mečkars. Les Rupane/Čergars II, qui se font appeler *Čergars* par tous les groupes de Roms, utilisent ce même nom pour désigner les Škodrànă/Čergars I. Cela ne signifie pas qu'ils ne délimitent pas bien leur groupe, mais que ce terme est devenu une simple étiquette[1]. *Čergar* a un jour fini par désigner l'« autre » Rom, celui d'un autre *fis*, ne parlant pas la même variété dialectale, jusqu'à être parfois repris par le groupe à son propre compte. On constate par exemple que les Arlis et les Bamils nomment *Čergars* à la fois les Škodrànă/Čergars I et les Rupane/Čergars II. Les seuls groupes qui semblent bien distincts sont donc les Mečkars (considérés comme les premiers arrivés dans le pays et les plus « albanisés ») et les Bamils (moins nombreux, dans moins de régions, et plus isolés des autres groupes de Roms).

Seuls deux groupes distinguent chaque groupe de Rom en l'appelant par un nom différent : les Mečkars et Škodrànă/Čergars I. Cela peut s'expliquer par une connaissance plus ample de leur communauté : il s'agit des groupes les plus anciennement implantés dans le pays en général et dans la capitale en particulier (celle-ci jouant un rôle de kaléidoscope ou de panorama de l'ensemble du pays). Les Arlis et les Rupane/Čergars II distinguent également tous les groupes par un nom différent, à l'exception des Rupane/Čergars II et des Škodrànă/Čergars I qu'ils nomment tous deux *Čergars*. Les Bamils appellent tous les Roms *Čergars*, à l'exception d'eux-mêmes (*Roms*) et des Mečkars qu'ils appellent *Mečkars*.

[1] Ce terme, omniprésent dans le Tableau 4, sert en effet à désigner, comme endonyme ou exonyme, à la fois les Škodrànă/Čergars I, les Rupane/Čergars II et les Arlis.

1.3.2. Mečkars

Le groupe des Mečkars[1] est considéré comme le plus ancien du pays – c'est-à-dire arrivé avant le XIXe siècle (voire depuis cinq siècles selon mon hôte). Selon Koinova (2000 : 14), qui s'appuie sur les données de Marcel Courthiade, les Mečkars se seraient d'abord implantés dans la plaine de Myzeqe et vivraient aujourd'hui principalement dans la plaine de Myzeqe, à Tiranë, Berat et Durrës. Ils constitueraient environ un tiers des Roms d'Albanie. En tout état de cause, ce groupe ne parle pas un dialecte arli, contrairement à l'affirmation de Friedman (2017 : note 10).

Ce groupe a été, selon Koinova (2000 : 6) et selon les dires de l'ensemble de nos informateurs, le premier à se sédentariser et à s'« albaniser ». Il bénéficie aujourd'hui de prestige et comporte une élite importante, tant sur le plan économique que culturel et social. C'est le groupe qui a développé le plus de contact avec les Albanais majoritaires, notamment par son travail dans le secteur agricole. Selon de De Soto, Beddies et Gedeshi (2005 : 9), les Mečkars travaillaient traditionnellement dans l'élevage et l'agriculture ; certaines familles possédaient des terres. La musique faisait partie des activités et sources de revenu traditionnelles à l'occasion des fêtes et des mariages, et ce encore aujourd'hui (Koinova 2000 : 14). Sous le régime de Hoxha, les Mečkars et d'autres Roms du sud du pays ont été embauchés dans les fermes et coopératives d'État, d'où les contacts étroits et nombreux avec les Non-Roms. Ils ont été les premiers à migrer, dès 1991, pour travailler à l'étranger, notamment en Grèce (De Soto, Beddies et Gedeshi 2005 : 13). Les enfants de ce groupe sont scolarisés en plus grand nombre et les jeunes mieux formés professionnellement. Ils habiteraient plus fréquemment que les autres dans des quartiers « mixtes » et seraient plus exogames que les autres Roms.

1.3.3. Arlis

Les Arlis sont un groupe depuis longtemps sédentaire, présent non seulement en Albanie, mais aussi au sud de la Serbie, au Kosovo, en Macédoine (Heinschink 1978) et en Grèce ; la plupart sont musulmans. Selon Heinschink (1978 : 12), le groupe des Arlis comprendrait en Yougoslavie plusieurs sous-groupes dont les noms dérivent de leurs métiers (*bugurdžije*, *konopljari*, *barudžije*, *arabadžije*) ou de leur région supposée d'origine (*Aškelije*, *Maljokja*, *Prištinelije*, *Topanelije*). Ils sont musulmans, à l'exception des chrétiens de Svetozarevo, aujourd'hui Jagodina. Friedman (2017) a récemment étudié les sept variétés différentes de l'arli qui coexistent à Skopje : Topaanli, Barutči, Madžur, Konopar, Prištevač, Gilanli, et Gavutno, présentant leur origine, leur époque d'arrivée dans la capitale, bouleversée par le tremblement de terre de 1963, ainsi que les principaux traits permettant de les distinguer.

[1] *Mečkar* signifie littérallement « oursier » (slave du sud *mečka*/мечка, 'ours' + suffixe agentif *-ar*). C'est l'exact équivalent d'*arixhi*, cité p. 29 (de l'albanais *ari*, ours et du suffixe agentif turc *-ci*). Ce groupe n'aurait aucun rapport avec les Mečkars de Yougoslavie et de Grèce (Koinova 2000 : 23, Friedman 2017 : note 10).

Les Arlis/Erlis[1] d'Albanie sont considérés comme un groupe arrivé récemment dans le pays, depuis centre trente ou cent quarante ans selon mon hôte. Ils seraient venus du nord-ouest de la Grèce, notamment la ville de Florina, d'où ils étaient arrivés depuis la Turquie continentale (Thrace) voire anatolienne si l'on en croit mon hôte (dont la grand-mère est originaire d'Izmir). Koinova (2000 : 14) confirme qu'ils seraient arrivés beaucoup plus tard que les Mečkars ; la migration se serait produite entre la fin du XIXe et les trois premières décennies du XXe siècle. Certain Roms de Korçë étaient musulmans et nommaient les chrétiens *das* (Mann 1933a) – terme que l'on retrouve également en Yougoslavie. Leur arrivée en Albanie est certainement due à « l'échange de population » arrangé entre les États grec et turc en 1924. Seuls les musulmans de « nationalité » albanaise, les Çams, avaient pu rester sur place, mais non les Roms et les Turcs. Certaines familles avaient choisi de migrer vers la Thrace orientale turque, tandis que d'autres avaient choisi de rester dans leur région en passant « simplement » la frontière albanaise : en Albanie se trouvaient beaucoup de Roms, de musulmans et de Roms musulmans. Hasluck (1938a : 54-55) écrit que « *Some who made Korcha their headquarters lived in Greek Macedonia until 1924, when they were due like other Mohammedan subjects of Greece to be sent to Asia Minor in exchange for Greek refugees; professional travellers though they were, they preferred the shorter journey to Albania.* » C'est l'histoire de la grand-mère de mon hôte, femme arli née à Florina (Grèce) en 1900, qui vint vivre en Albanie où elle eut quatre enfants à Korçë entre 1924 et 1938. Notons que le terme pour chrétien n'est plus, aujourd'hui à Korçë, *das*, mais *gadžo*, littéralement « non-rom ». Ceci est cohérent avec ce qu'affirme Heinschink (1978 : 12) des « tribus arlisques » de Yougoslavie : « Eux-mêmes préfèrent s'appeler '*Khorane*' ou '*korahane roma*', donc 'Tsiganes turcs', pour se différencier des autres groupes qui n'ont pas, comme eux, subi la domination et l'influence culturelle turco-islamique et qui sont pour eux des 'gadžikane roma' ('Tsiganes paysans ou chrétiens'). La famille est de tradition musulmane et s'est officiellement convertie à l'orthodoxe en 2016, devenant *gaxhe* selon ses propres termes. D'autres familles, enfin, choisirent de se convertir à l'orthodoxie afin de rester en Grèce, migrant si besoin vers le sud : c'est le cas des Roms de Parakalamos (à mi-chemin entre Ioannina et l'Albanie) dont Matras (2004 : 60-61) décrit la variété. Selon lui, des liens familiaux ont été conservés entre les familles des villages (convertis) à la frontière avec l'Albanie, les familles (musulmanes) ayant migré en Albanie même et les familles (converties) ayant migré à Parakalamos.

Selon Koinova (2000 : 14, 23), qui s'appuie sur les données de Marcel Courthiade, les Arlis vivraient aujourd'hui principalement à Korçë, Tiranë, Elbasan et Berat. Leur importance numérique est au moins similaire à celle des Mečkars. Cependant ce groupe compte plus de personnes pauvres, avec un grand nombre d'enfants non scolarisés voire non-enregistrés à l'état-civil. Il compte également une élite économique et culturelle, même si elle semble moins importante que celle des Mečkars. Sous le régime, les Arlis ont exercé différentes professions, mais plutôt dans le domaine « privé » que dans le domaine public comme les Mečkars. Ils pratiquaient essentiellement l'artisanat

[1] *Arli/erli* signifie « autochtone, local, établi » (turc *yerli*, local, de *yer*, endroit + suffixe d'origine *-li*) est emprunté au turc *erli/yerli*, par opposition à de nouveaux arrivants et aux nomades (qui étaient/sont loin d'être tous Roms dans les Balkans). Selon Friedman (2017 : 33), la perte du yod initial correspond au traitement macédo-bulgare du *ye-* initial du turc, tandis que le passage de *-e-* à *-a-* à l'initiale reflète un phénomène du grec moderne et du macédonien pour les mots turcs. Selon Heinschink (1978 : 12), « cette dénomination correspond bien au mode de vie des tribus arlisques qui ont été sédentarisées depuis longtemps ». Friedman (2017 : 33) confirme qu'en Macédoine, la plupart des Arlis étaient des paysans, à l'exception de ceux établis dans les zones urbaines, qui pratiquaient une grande variété de professions au XIXème siècle : à Skopje ils pouvaient être tanneurs, musiciens, docteurs, avocats.

et le commerce : maquignonnage, chaudronnerie, vannerie, vente de vêtements d'occasion (De Soto, Beddies et Gedeshi 2005 : 11-13). Comme les Mečkars, ils étaient des musiciens reconnus et appréciés. Les Arlis travaillant dans le secteur public sont restés, comme les Čergars et contrairement aux Mečkars, confinés à des emplois sous-qualifiés. C'est pourquoi la plupart s'est tournée vers le secteur privé informel (De Soto, Beddies et Gedeshi 2005 : 12). Leur activité s'est cependant trouvée en déclin après la chute du régime : ils ont dû, comme beaucoup d'autres, migrer à l'étranger, notamment en Grèce et en Italie. Ils y ont travaillé de manière permanente ou saisonnière (De Soto, Beddies et Gedeshi 2005 : 14), dans l'agriculture ou le commerce – du moins jusqu'à la récente crise économique.

1.3.4. Čergars I et II

Les Čergars sont considérés comme un groupe arrivé récemment dans le pays – à la fin du XIXe et au début du XXème siècle. Selon mon hôte, ils proviendraient, comme les Arlis, de Turquie et auraient migré à la même époque période. La variété des Čergars que j'ai pu rencontrer dans la famille de son épouse présente cependant des romanismes et des slavismes plutôt que des turcismes et des hellénismes. Il s'agit donc de deux groupes différents de « Čergars » (cf. section 1.3.1, p. 17).

Selon Marcel Courthiade cité par Koinova (2000 : 14), le premier groupe de Čergars provient des principautés roumaines, aurait traversé la Serbie et le Kosovo avant de s'implanter à Shkodër en Albanie (Koinova 2000 : 14). Il s'agirait du même groupe que celui des Gurbets de Yougoslavie (Kosovo, Monténégro, Serbie), avec lesquels ils entretiendraient encore des contacts (Koinova 2000 : 23). Selon Friedman (2017 : note 10), le groupe des Čergars (du turc, via le slave, pour « habitant de tentes ») est le même que le groupe Gurbet (du turc et du slave via le turc pour « migrant »), Džambaz (du turc pour « maquignon »), Vlah (du slave pour « valaque/roumain »), Leaš (du roumain pour « nomade »), Kalderaš (du roumain pour « chaudronnier »). Stuart Mann a rencontré des représentants de ce groupe qu'il nomme les Scutariens. Les hommes chassaient, tuaient des chiens errants (activité à l'époque rémunérée) et pratiquaient parfois le commerce de chevaux. Selon lui, ces Roms viendraient de la région de « Gegënía » (Gegëria) et de la Serbie (Mann 1933a : 2). Ce groupe vit aujourd'hui principalement dans la partie centrale de l'Albanie, à Tiranë ou encore à Durrës.

L'ethnologue Weigand mentionne les Tsiganes de Macédoine qu'il a pu observer dans ses enquêtes de terrain en 1917-1918. Rappelons que la Macédoine de son époque comprend les régions de Skopje, Bitola, Ohrid, mais aussi Serrès, Thessalonique, Veria, Naoussa et Edessa, soit les parties slavophone et hellénophone de la Macédoine dite « historique ». Korçë se trouve dans la partie albanophone de celle-ci.

> « *Die Gesamtzahl mag zwischen 60—80 liegen, wovon etwa zwei Dritten Mohammedaner sind. Die übrigen sind orthodoxe Zigeuner, die neben ihrer Muttersprache auch bulgarische sprechen. […] Ihre alte Sprache haben sie bewahrt, doch haben sie viele türkische und bulgarische Elemente eingemischt. Neugriechische sind Gemeingut aller Zigeuner. Alle sprechen die Sprache ihrer Umgebung mit Geläufigkeit, aber doch mit charakteristischer Aussprache. Zuweilen habe ich auch rumänisch sprechende Zigeuner getroffen, und zwar umherziehende Bärenführer.* » (Weigand 1924 : 86)

Il existe donc, tout près de Korçë, de nombreux Roms musulmans, parlant probablement le turc, ainsi que des Roms orthodoxes parlant le bulgare et des Roms nomades parlant le roumain[1]. Les Čergars I pourraient être issus du deuxième ou du troisième groupe mentionné. Les Čergars II pourraient être issus du premier groupe mentionné. Tous ont en commun de parler le romani avec des emprunts au turc, au bulgare et au grec. D'autres sources historiques le confirment : ainsi Wallisch (1931 : 76) parle-t-il de Tsiganes nomades arrivant de Serbie du sud en Albanie. Ceux-ci sont à distinguer des Tsiganes sédentaires, qui sont artisans ou petits paysans, et dont les femmes sont domestiques[2] en ville et appréciées – il s'agit là des Mečkars et des Arlis.

Le second provient de la Turquie d'Europe, aurait traversé la Macédoine slave avant de venir en Albanie par le sud (Përmet) : Peqin puis Tiranë. Mon hôte, en racontant l'histoire de ce groupe, croyait me parler de l'histoire de sa belle-famille, alors que l'histoire de cette dernière est probablement différente. En effet, sa belle-famille appartient aux Čergars I de la zone roumanophone et slavophone.[3] En revanche, il s'agit de l'histoire de son grand-père, de père et de mère čergar, qui se maria successivement à deux femmes arli et vint vivre à Korçë. Le destin de ce groupe est le même que celui des Arlis à cette époque. Il est possible que Stuart Mann ait rencontré des représentants de ce groupe parmi les Korçiens, notamment Neğipi, qui aurait quitté la Macédoine grecque pour éviter l'échange de population avec la Turquie. Celui-ci voyagea, avec sa femme, à travers toute l'Europe et aux États-Unis où il travailla comme clarinettiste, avant de rentrer en Albanie. Je n'ai pu rencontrer de locuteur parlant cette variété dialectale.

Les deux groupes ne sont cependant pas dissociés dans les rares études sur les Roms albanais. On y apprend que les « Čergars » sont restreints du point de vue numérique, mais disposent d'une élite économique et culturelle importante. Ils pratiquaient traditionnellement le maquignonnage, la chaudronnerie, la vannerie et la vente de vêtements d'occasion (De Soto, Beddies et Gedeshi 2005 : 11, 13). Il semble qu'ils aient eu une tradition de vente ambulante et de nomadisme (Koinova 2000 : 6, qui cite Marcel Courthiade). Sous le régime, les Čergars ont exercé différentes professions, souvent sous-qualifiées et essentiellement dans le secteur informel, à l'instar des Arlis (De Soto, Beddies et Gedeshi 2005 : 12). Après la chute du régime hoxhiste, beaucoup ont dû migrer à l'étranger, de manière permanente ou temporaire, pour trouver du travail (De Soto, Beddies et Gedeshi 2005 : 14).

1.3.5. Bamils

Il fut difficile d'obtenir des informations sur le groupe des Bam(b)ils/Bamirs[4] car il est mal connu des autres Roms. J'ai rencontré quelques personnes ayant un parent ou une partie de la famille issue

[1] Il pourrait s'agir également de Beaşi/Rudari, groupe nomade souvent assimilé aux Roms, parlant une variété de roumain mais non le romani.

[2] Fait également mentionné par Hasluck (1938a).

[3] S'il y a erreur, elle est peut-être due à l'homonymie des deux groupes.

[4] Nous ignorons l'étymologie de ce mot.

de ce groupe, mais elles ne m'en ont pas parlé, et pratiquaient de manière prédominante la variété de l'autre partie de leur famille, voire l'albanais. Selon mon hôte, ils auraient migré au XXe siècle en même temps que les Čams[1] et parleraient encore grec. Marcel Courthiade confirme que ce groupe serait présent en Albanie depuis moins d'un siècle et, pour beaucoup, depuis les années 1950. Ils semblent qu'ils soient peu nombreux et ne soient présents qu'à Fier et dans le sud du pays (Gjirokastër, Delvinë).

Comme les Arli et les Čergars, ils auraient bénéficié du droit, sous le régime de Hoxha, de poursuivre leurs activités traditionnelles, même si celles-ci incluaient la propriété privée (De Soto, Beddies et Gedeshi 2005 : 11). Leur profession traditionnelle serait l'étamage et la confection de torréfacteurs à café, de marmites et de casseroles (De Soto, Beddies et Gedeshi 2005 : 10). Ce groupe et celui des Kurtofis comprendraient les Roms les plus pauvres du pays, et vivraient pour beaucoup dans des ghettos voire des bidonvilles.

1.3.6. Kurtofis

Il ne m'a pas été donné non plus de rencontrer de Roms du groupe des Kurtofis/Kortufis[2], considéré comme le groupe le plus restreint du point de vue numérique. Ils vivraient dans de petits villages dans le sud et jusqu'à Fier (Koinova 2000 : 14). Avec les Bamils, il s'agirait des Roms les plus pauvres du pays, avec un grand nombre d'enfants non scolarisés, non-enregistrés à l'état-civil, vivant dans des bidonvilles. Selon un informateur arli (cf. exemple (186), corpus p. 568), ils vivraient dans des caravanes et seraient attachés à la tradition. Koinova (2000 : 6, 11) confirme que ce groupe pratique encore aujourd'hui le commerce de proximité et l'artisanat ; ils porteraient encore aujourd'hui des vêtements traditionnels colorés et les femmes porteraient un pantalon sous une longue robe, ainsi qu'un foulard sur la tête.

1.3.7. Différences entre les groupes

Les groupes se distinguent par la variété qu'ils parlent, par leurs métiers traditionnels, mais aussi par leurs coutumes si l'on en croit (Koinova 2000 : 14). Une étude ethnologique approfondie serait cependant bienvenue pour comprendre en quoi consistent ces différences. Celle menée par De Soto, Beddies et Gedeshi (2005) permet d'observer des différences sur le plan sociologique, notamment le niveau d'instruction et les professions exercées. Le Tableau 5 confirme que l'illettrisme est, parmi les jeunes interviewés, le plus faible chez les Mečkars (38,8 %), qui se distinguent nettement des Čergars (63,3 %), Arlis (désignés par l'exonyme *Karbuxh*[3] : 77,9 %) et Bamils (86,7 %) – dans cette enquête également, le manque de données quant aux Kurtofis (100 %) se fait sentir. Ces taux étaient

[1] Albanais du nord de la Grèce, expulsés par les autorités, qui ont migré au sud de l'Albanie à la fin de la Seconde Guerre Mondiale.

[2] Nous ignorons l'étymologie de ce mot.

[3] Variante de kabudži/kalbudži, nom donné par les Mečkars aux Arlis (cf. Tableau 4 p. 11).

nettement moins élevés sous le régime (les Mečkars ne comptent que 10,8 % d'illettrés parmi les 21-40 ans), mais la hiérarchie entre groupes était la même.

Tableau 5 : Taux d'alphabétisation des Roms par groupes (De Soto, Beddies et Gedeshi 2005 : 189)

Table D.3. Illiteracy Rate and School Years Completed, by Age Group and Roma *Fis*

	Age group 7–20			Age group 21–40		
Roma *Fise*	Illiteracy (absolute)	Illiteracy (percentage)	Number of school years completed	Illiteracy (absolute)	Illiteracy (percentage)	Number of school years completed
Meçkar	80	38.8	4.46	22	10.8	6.85
Karbuxh	286	77.9	3.54	164	57.7	5.47
Cergar	38	63.3	3.18	15	51.7	5.36
Bamill	26	86.7	1.25	16	66.7	4.38
Kurtof	1	100	—	1	100	—
Total	**431**	**62.2**	**4.02**	**235**	**39.6**	**6.22**

Source: Socioeconomic household questionnaires with Roma and Egyptians for the Needs Assessment (2002/03).

1.4. Roms et Égyptiens

Il existe en Albanie une minorité différente des Roms, mais qui partage en grande partie la même histoire de pauvreté, d'exclusion et de discrimination. Il s'agit des « Égyptiens » (*Jevg/Jevga, Jevgjit/Jevgitka, Evgjit/Evgjitka*), certainement apparentés au groupe des Balkano-Égyptiens qui vivent également en Macédoine, au Kosovo, au Monténégro et en Serbie – où ils se nomment *Ashkali·je*[1]. Ils ont pour langue l'albanais et non le romani. Les Albanais majoritaires, notamment dans la partie nord du pays et au Kosovo, utilisent parfois un terme chapeau pour désigner les Roms et les Égyptiens, *Magjup·e, Magjyp·e, Maxhup·e, Xhip/Xhipka*, équivalent du français « Tsigane ». Dans la région de Shkodër, ce terme désignerait à l'origine exclusivement les Égyptiens (Courthiade 1988b). Il a la même étymologie qu'« Egypte » et que l'arabe *Kibtian*, assorti selon Hasluck (1938a : 51) d'un obscur préfixe albanais *ma-*. Friedman (2017 : note 8) mentionne un groupe de forgerons du sud-ouest de la Macédoine parlant l'albanais et non le romani, dont l'endonyme est *Gjupci*, qu'il traduit par *Gypsies*, mais qui ne s'auto-identifient pas pas comme Roms. Il précise que dans le contexte de politisation des identités ethniques à partir de la fin du XXème siècle, ils se sont fait appeler *Egipkjani*, Égyptiens. Il s'agit fort probablement du même groupe.

Le terme *Jevg* est souvent considéré, à tort, par les dictionnaires, comme un équivalent non péjoratif (donc politiquement correct) de *Magjup*, donc un terme-chapeau pour les deux communautés, alors qu'il désigne en réalité spécifiquement l'égyptienne (Koinova 2000 : 13). Le terme aujourd'hui utilisé dans les cercles associatifs ou en contexte militant est le terme *Egjiptjan/*

[1] Piasere (1982) écrit cependant de ce groupe qu'il s'agit de « "sang-mêlé", moitié rom moitié albanais ».

Egjiptjanka. Notons qu'à l'époque de Mann (1933a : 2), les Égyptiens auraient été nommés *Bičəs* dans le sud de l'Albanie, mais Hasluck (1938a : 51) ne trouve pas trace de cette appellation. Selon Mann (1933a : 2) et malgré ce qu'affirmera plus tard (Newman 1969), les Albanais majoritaires distinguent très bien les deux communautés – et c'est encore le cas aujourd'hui.

Lors du recensement de 2011, 3 368 personnes se seraient déclarées égyptiennes, soit 0,12 % de la population totale et presque deux fois moins que les Roms. Mais les Égyptiens, puisqu'ils se distinguent du reste des Albanais surtout par des critères socio-économiques, se font en général recenser dans la catégorie majoritaire, d'autant qu'ils déclarent l'albanais pour langue maternelle. Notons que Koinova (2000 : 13) estime, avec Marcel Courthiade, à 10 000 les Égyptiens vivant dans la seule ville de Tiranë.

Les origines des « Égyptiens » sont obscures. Certains chercheurs pensent qu'elles sont les mêmes que celles des Roms. La minorité égyptienne serait simplement issue d'un groupe de Roms qui, pour monter dans l'échelle sociale, se seraient assimilés linguistiquement et auraient créé une identité séparée de leur groupe d'origine (c.p. de Carol Silverman en 2001, citée par De Soto, Beddies et Gedeshi 2005 : 148). D'autres chercheurs pensent au contraire que leurs origines sont différentes. Les preuves et les sources historiques manquent, mais non les hypothèses : cf. Wallisch (1931 : 76), Mann (1933a : 3), Hasluck (1938a : 49-50), Newman (1969), Courthiade (1988b), Koinova (2000 : 4), De Soto, Beddies et Gedeshi (2005 : 7-8), Marushiakova et Popov (2016 : 37-39)…

Le premier élément distinguant les Égyptiens est qu'ils ne parlent pas romani et ne l'ont, selon les témoignages remontant jusqu'au début du XXème siècle et la mémoire des aïeux de nos informateurs, jamais parlé – ce qui était déjà le cas dans les années 1930[1]. Ils parlent albanais avec un accent particulier si l'on en croit Mann (1933a : 3), quelques mots spécifiques à leur groupe et revêtant une fonction de secret ou d'exclusion (Mann 1933a : 3-4, Koinova 2000 : 24). Ces mots sont soit des mots albanais dont le sens est détourné, soit des emprunts au turc, à l'italien ou au romani. On ne peut donc pas en conclure que les Égyptiens ont une langue propre[2]. Selon (Mann 1933a : 4), beaucoup d'Égyptiens parlent italien, mais selon (Hasluck 1938a : 54), seuls ceux de Tiranë et de Durrës le parlent, en revanche les anciens parlent encore le turc. Le deuxième élément est qu'ils semblent avoir été sédentaires et jamais itinérants, ce qui n'est pas le cas de tous les groupes de Roms albanais. Le troisième élément est leur propre conscience identitaire, distincte de celle des Roms (déjà affirmée par Mann 1933a : 2).

Si l'on en croit Mann (1933a : 3) et Hasluck (1938a : 53), les relations entre les deux groupes sont empreintes de distance et de mépris, et il y a peu d'intermariages. Selon Hasluck, les Tsiganes nomades ne donnent jamais leurs filles en mariage aux sédentaires, mais leur « volent » parfois une fille. Ceci est confirmé par une étude américaine citée par Koinova (2000 : 13), selon laquelle les Égyptiens pratiqueraient en revanche l'exogamie avec les Albanais majoritaires. Les pratiques de mariage sont donc différentes – par exemple les remariages sont moins importants chez les Roms que

[1] Si les Égyptiens sont des Roms ayant cessé de transmettre la langue, l'événement s'est donc produit au plus tard au courant du XIXème siècle.

[2] Ce qui n'empêche pas certains activistes, pour renouer avec les origines présupposées par leur ethnonyme, de prendre contact avec l'État d'Egypte et de militer pour l'apprentissage de l'égyptien ancien.

chez les Égyptiennes, mais plus importants que chez les Albanais majoritaires (Koinova 2000 : 10). Les Roms sont traditionnellement endogames, mais pratiquant l'exogamie d'un groupe de Roms à l'autre : cette exclusion traditionnelle[1] des Égyptiens pourrait constituer un quatrième élément de distinction entre les deux communautés, qu'elles aient ou non été un même groupe ethnique à l'origine. Selon nos informateurs roms, les Égyptiens ne se considèrent pas tous comme ethniquement distincts de la majorité albanaise – si ce n'est socio-économiquement. C'est que, comme les Roms, les Égyptiens ne se distinguent pas par la religion mais que, contrairement aux Roms, ils ne se distinguent pas non plus par la langue. L'étude de De Soto, Beddies et Gedeshi (2005) montre d'ailleurs que Roms et Égyptiens ne se démarquent pas de la majorité albanaise par les mêmes éléments culturels. Sula Raxhimi (2015 : 42) choisit d'ailleurs de ne pas considérer, elle non plus, les Roms et les (Balkano-) Égyptiens comme un même groupe anthropologique que les Roms.

Enfin, les Roms nomment les Égyptiens *Čhibane*, parleurs, bavards,[2] et les considèrent comme un peuple à la fois différent des Roms (puisque, par exemple, ils ne parlent pas romani) et de la majorité albanaise (puisque, par exemple, ils sont très ou plus sombres de peau). Selon Koinova (2000 : 13) qui cite Marcel Courthiade, les Roms considéreraient les Égyptiens comme des *gadže*, donc des Non-Roms, notamment parce qu'ils ont l'albanais pour langue maternelle et parce qu'ils pratiqueraient l'exogamie. Cependant l'un de nos informateurs, Arli de 70 ans, nous a expliqué (cf. interview du corpus spontané, p. 577) que les Égyptiens se nomment *čhibane Roma*, Roms bavards, c'est-à-dire qu'ils sont des Roms qui ne savent pas la langue. À la question de pourquoi ils ne savent pas la langue, l'informateur répond : parce qu'ils ne sont pas Roms. On voit à quel point, selon lui, la connaissance de la langue est fondamentale dans l'identité rom.[3] Ceci est confirmé pour d'autres régions. Cela pourrait appuyer la théorie selon laquelle les Égyptiens serait un groupe assimilé à la majorité albanophone : à partir du moment où ils ont renoncé à la langue, ils ont également renoncé à être considérés voire à se considérer comme Roms (Marushiakova et Popov 2016 : 37-39). Les Égyptiens semblent donc constituer une catégorie ethnique de transition entre les Roms proprement dits et les *gadže* proprement dits. Des stéréotypes courent parmi les Roms à propos des Égyptiens, qui rappellent en fait ceux qui courent parmi les Albanais majoritaires à propos des Roms : les Égyptiens seraient par exemple plus riches qu'ils ne le laissent croire, certains étant millionnaires[4]. Si Roms et Égyptiens sont conscients d'être deux groupes différents, et veulent souvent l'être, et si la plupart des Albanais majoritaires qui les côtoient en sont également conscients, d'autres, qui ne fréquentent donc ne connaissent ni les uns ni les autres, utilisent le même ethnonyme pour les désigner, négligeant leurs différences culturelles. L'unité de ces deux groupes réside alors dans le regard porté de l'extérieur, les difficultés socio-économiques et les discriminations dont ils souffrent (De Soto, Beddies et Gedeshi 2005 : 10). Comme les « Tsiganes » ou aujourd'hui les « Roms » en France, les Égyptiens et les Roms entrent alors dans une même catégorie sociale plutôt qu'ethnique.

[1] J'ai tout de même rencontré sur le terrain des personnes nées de parents mixtes, Roms et Égyptiens, ainsi que des Roms mariés à des Égyptiens.

[2] Selon Koinova (2000 : 13), les Roms nommeraient ironiquement les Égyptiens *sir*, ail, ce que je n'ai pu vérifier.

[3] C'est pourquoi la langue est un instrument essentiel dans le militantisme rom, cf. Matras (2013a) et Garo (2002).

[4] L'un des millionnaires impliqués dans la crise des Pyramides de 1997 était un Égyptien (Koinova 2000 : 13).

2. La langue romani

Le romani est une langue indo-européenne de la branche indo-iranienne, du groupe indien central. Elle comprend, comme toute langue, de nombreuses variétés qui s'expliquent par la dispersion ancienne des Roms sur différents continents, par leur division en petites unités ainsi que par les diverses langues de contact (qui varient dans le temps et dans l'espace).

Les extraits en langue romani qui ne font pas partie de notre propre corpus sont cités en respectant l'orthographe originale de l'auteur qui les a publiés.

2.1. Dialectologie

2.1.1. Spécificités de la langue

Un premier problème est que l'on dispose de peu d'attestations historiques du romani et donc de peu de perspective diachronique satisfaisante. On dispose toutefois de nombreuses listes de mots et de quelques extraits de textes, mais non d'un échantillon de toutes les variétés. Les premiers fragments de texte que l'on connaît en Europe occidentale datent des environs de 1515 (Munich), avec soixante mots latins traduits en romani, et de 1547 (Pays de Galles / Montpellier), avec quinze phrases en romani (Courthiade 2014a : 1). Il existe également une lettre écrite pendant la Guerre de Trente Ans, publiée par Courthiade (2014a : 9). En Europe orientale, on dispose d'un texte transcrit autour de 1783 par Marsden en dialecte de Roumélie, puis de transcriptions de Paspati (1870). Friedman et Dankoff (1991) ont publié et analysé un texte plus ancien encore : une liste de mots et de phrases notés en 1668 en Thrace à Komotini (Gümülcine/ Κομοτηνή/Гюмюрджина), par le célèbre voyageur et écrivain ottoman Evliya Çelebi dans le Seyāḥat-nāme, *Livre des voyages* (Çelebi 1630). Il s'agit donc de la variété de Thrace, au carrefour des langues et de la diversité culturelle.

Un deuxième problème est culturel. On sait également la circonspection et la défiance des Roms envers les Non-Roms, qui fait que « fort polyglottes, ils parlent rarement leur langue propre devant les étrangers » (Bourcart 1922 : 276). Cela amène nombre d'observateurs et de chercheurs à croire que tel ou tel groupe ne parle pas ou plus romani.

À la rareté des attestations vient s'ajouter le problème de la dispersion géographique dans différents pays et leur présence territoriale discrète dans les pays en question. L'exode de l'Inde est l'élément fondateur du romani en tant que langue et la raison de la présence pluriséculaire de cette langue indienne en Europe (puisqu'il ne s'agissait, auparavant, pas de romani mais d'une langue indienne centrale). La migration est la réalité de nombreux groupes, à de nombreuses époques, de manière temporaire (nomadisme ou semi-nomadisme) ou constante. Les déplacements mettent en contact les variétés entre elles, brouillant les cartes de la généalogie des variétés, mais également les variétés avec d'autres langues. Enfin, les Roms ne sont la population majoritaire d'aucune zone plus vaste qu'un quartier ou qu'une commune (c'est le cas de Šuto Orizari en Macédoine).

La mobilité est un autre élément fondamental, donc le contact de langues, en raison non seulement des migrations, mais également du caractère partout minoritaire des Roms. Matras (2002 : 191) parle d'une langue en permanence en contact (cf. section 2.2, p. 34). On a déjà remarqué, en section 1.3.1, p. 17, que les noms des différents groupes étaient empruntés à des langues de contact. C'est le cas de beaucoup de groupes en Europe, et de quasiment tous ceux des Balkans, le plus souvent des noms de profession ou de statut social (Matras 2004 : 67-68).

Les dialectologues s'accordent sur trois points. D'une part, les différentes variétés parlées n'empêchent pas de considérer le romani comme une langue unique, « malgré » sa diversité. D'autre part, la classification ne se fait pas seulement par la zone géographique, mais par l'histoire migratoire du groupe, le contact avec d'autres variétés ou d'autres langues, etc. Au contraire des langues territoriales, pour lesquelles on peut faire une dialectologie « aréale » (et pour laquelle on peut parler de *Flächendialekte*), le romani, est une langue diasporique, pour laquelle on peut faire une dialectologie « insulaire » (et pour laquelle on peut parler d'*Inseldialekte*) (Boretzky 1998). Enfin, les variétés portent le plus souvent le nom des groupes (dans le sens de clan ou de *fis*, cf. section 1.3.1 p. 17) qui les parlent. On peut tracer un trait d'égalité entre le groupe (par exemple : les Mečkars) et la variété parlée par ce groupe (le mečkar). Ceci n'empêche pas les cas d'homonymie : il existe en Albanie deux groupes appelés *Čergars*, les *Mečkars* d'Albanie ne sont pas les *Mečkars* de Yougoslavie et les *Ka(l)budžija* (Arlis) albanais sont différents des *Kalburdjia* bulgares décrits par Gilliat-Smith (1915 : 6).

2.1.2. Modèles dialectologiques

Les études dialectologiques, si elles sont issues de la linguistique historique et comparée, présentent une approche diachronique, non via une généalogie que l'on ne peut retracer, mais via les emprunts aux langues étrangères considérés comme témoins du trajet migratoire des groupes (dès Miklosich 1872). Matras (2013b : 203) analyse les présupposés de cette démarche, notamment le fait que l'installation d'un groupe de Roms dans une région donnée aille de pair avec une influence culturelle et linguistique, et donc une distinction croissante du groupe en question par rapport aux autres groupes de Roms. Des arbres représentant les dialectes et groupes de dialectes peuvent alors être dessinés, non sur la base de changements internes à la langue, mais sur celle de l'apport extérieur du contact (Matras 2005a : 9) – ce qui est une forme de détournement du modèle phylogénétique.

Gilliat-Smith (1915), dans son étude sur le romani du nord-est de la Bulgarie, compléta cette démarche en s'appuyant également sur les innovations internes à la langue et sur les mouvements de population, plutôt que la stricte géographie des dialectes. Il introduisit une distinction entre les variétés des nomades pour la plupart chrétiens (qu'il nomme par leur exonyme : *Vlachs*), influencées par le bulgare et le roumain, et celles des sédentaires pour la plupart musulmans (qu'il nomme *Non-Vlachs*) influencées par le turc. Il considère cette distinction comme de nature génétique, expliquant la présence des dialectes *vlach* comme le produit de l'immigration de ces groupes à l'origine roumanophones en zone bulgare. Depuis l'abolition de l'esclavage dans les principautés roumaines au milieu du XIXe siècle et jusqu'à nos jours, l'émigration massive et répétée de ces groupes en fait les plus dispersés, les plus visibles et présents partout en Europe. La distinction *vlax/non-vlax*, ainsi

orthographiée, est devenue, depuis Gilliat-Smith (1915), fondamentale en dialectologie et sert de repère de classification quelle que soit la variété et la région d'étude (Matras 2005a : 9, Matras 2013b : 203). Lorsque Uhlik (1956) opère par exemple une classification des parlers de Yougoslavie, il commence par distinguer les parlers *vlax* des parlers non-*vlax*. C'est en grande partie la même que l'on trouve chez Courthiade (2009a)[1] entre deux super-dialectes, baptisés E (originaire de Valachie et de Moldavie, dispersé partout en Europe depuis le XIXe siècle) et O (originaire des Balkans, dispersé partout en Europe depuis l'arrivée des Roms en Europe). Selon Matras (2013b : 204), cette classification donnerait une préséance trop importante aux innovations du groupe *vlax*/E, au détriment de celles du groupe non-*vlax*/O, dont beaucoup sont plus anciennes. Des approches alternatives ont été proposées, par exemple Vencel' et Čerenkov (1976), que l'on peut considérer comme l'ancêtre de la classification consensuelle actuelle (Matras et Bakker 1997, Matras 2005a), où les dialectes dits *vlax* constituent un groupe parmi d'autres. Friedman et Dankoff (1991), quant à eux, remettent en question l'importance fondamentale qu'a prise la distinction *vlax*/non-*vlax* dans la distinction des dialectes des Balkans.

Plusieurs recherches ont été réalisées dans le domaine de la génétique sur les populations pour venir étayer la classification dialectale phylo-génétique (Bakker 2012). Elles tentent notamment de reconstruire le patrimoine génétique du plus ancien ancêtre commun de la population rom, dont la datation varie grandement, ainsi que l'époque où les groupes de population ont commencé à se séparer. Les études concluent que l'endogamie est très ancienne chez les Roms, mais que la population est génétiquement très hétérogène, et ce plus que les populations européennes autochtones (Bakker 2012 : 104). Cette hétérogénéité n'a fait que s'accentuer avec la dispersion et la fragmentation des Roms en petits groupes. La diversité génétique n'est pas corrélée à la géographie ou au pays de résidence, mais aux groupes dialectaux et à leur histoire migratoire (Bakker 2012 : 105, citant Gresham et al. 2001 : 1328).

L'approche la plus récente est le modèle des vagues, initié par Matras (2002 : chap. 9) et détaillé dans Matras (2005a). Elle prend en compte à la fois les traits hérités, acquis par changement interne, acquis par contact, la position géographique des groupes et les migrations passées. Elle conditionne la position d'un dialecte dans la classification à sa position géographique au moment où se diffuse une innovation structurelle. Selon l'auteur, les variétés sont reliées par des innovations internes se diffusant d'un groupe à l'autre – ce qui n'est pas sans rappeler la *Wellentheorie*, théorie des vagues de Schmidt (1872). On peut donc les trouver partagées par des populations voisines, géographiquement contiguës et/ou en contact régulier à l'époque de naissance et de diffusion de l'innovation. Deux dialectes ne peuvent partager un trait que s'ils se trouvent du même côté de la ligne marquant l'extension de la zone de diffusion de ce trait (Matras 2005a : 30-31) : l'affinité entre les dialectes consiste en l'adoption d'un ensemble de traits similaires. Elle n'est donc pas génétique mais en quelque sorte adoptive. Le romani forme donc bel et bien un continuum de dialectes reflétant l'histoire de ses innovations structurelles (lorsqu'elles ne sont pas contingentes), de ses conservatismes (lorsqu'ils présentent une régularité) et de ses locuteurs (contact). Il s'agit donc d'un modèle dynamique, considérant les traits structuraux comme le produit d'un processus graduel d'exposition à une innovation, d'accommodation puis d'adoption ou de rejet du trait en question

[1] L'autre distinction qu'il opère, entre dialectes à mutation *versus* sans mutation consonantique n'a en revanche pas son équivalent ailleurs. Elle concerne les deux superdialectes, mais surtout le E.

(Matras 2010 : 47). Les études permettent de constater l'existence de grands espaces de diffusion cohérents, qui permettent de classer les variétés par groupes. Ces espaces présentent en général un centre (plus innovant) et une périphérie (plus conservatrice) qui ne peuvent pas s'expliquer par un modèle génétique attribuant les innovations aux migrations aléatoires de groupes parlant des variétés distinctes et préexistantes.

2.1.3. Classification dialectologique

Pour une présentation générale de la dialectologie romani, on pourra lire Matras (2002 : 222-225) et Boretzky (2007). La plupart des linguistes s'accorde sur le regroupement suivant Matras (2005a : 9-10 et Matras (2010 : chap. 2), représenté par la Figure 1 :

- **Groupe *vlax*** : variétés originaires de Valachie et de Transylvanie, parlées par des groupes initialement nomades[1]. Les locuteurs de ces variétés sont très nombreux et aujourd'hui dispersés non seulement en Europe mais aussi en Amérique (par exemple au Mexique : cf. Adamou 2013 et Adamou 2016). On peut opérer une subdivision entre variétés méridionales (gurbet, džambazi...) et septentrionales (kalderaš, lovari, čurari...), parfois nommées respectivement ancien *vlax* et nouveau *vlax* (Bakker 2012 : 102).
- **Groupe central** : variétés d'Europe centrale, que l'on peut subdiviser entre variétés septentrionales (sud de la Pologne, nord de la Slovaquie, Ukraine transcarpathique), dont le bergitka, et méridionales (sud de la Slovaquie, Hongrie, sud-est de l'Autriche, nord de la Slovénie), dont le romungro ou le vend.
- **Groupe des Balkans** : variétés du centre et du sud de la péninsule balkanique, ainsi que du pourtour de la Mer Noire, subdivisées en Balkans méridionaux I et Balkans méridionaux II (cf. section 2.3, p. 37).
- **Groupe du nord** : ce groupe inclut géographiquement l'ouest et le nord de l'Europe, mais aussi les péninsules italienne et ibérique (Bakker 1999). Matras (2002) préfère considérer un groupe du **nord-est** (zone baltique et russe) et un groupe du **nord-ouest** (zone germanique et scandinave), dont le sinto/manouche ou le kaale.
- **Isolats** : romani des Abruzzes, de Calabre, de Slovénie, d'Azerbaïdjan, romani gallois†, errumantxela†…

[1] Cette appellation, due à Gilliat-Smith (1915), n'est en réalité pas très heureuse, si l'on considère la polysémie du terme *vlax* : il désigne quasiment partout les populations (initialement nomades ou semi-nomades) parlant une langue romane, dispersées dans les Balkans (Aroumains, Mégléno-Roumains, Beaşi/Rudari…).

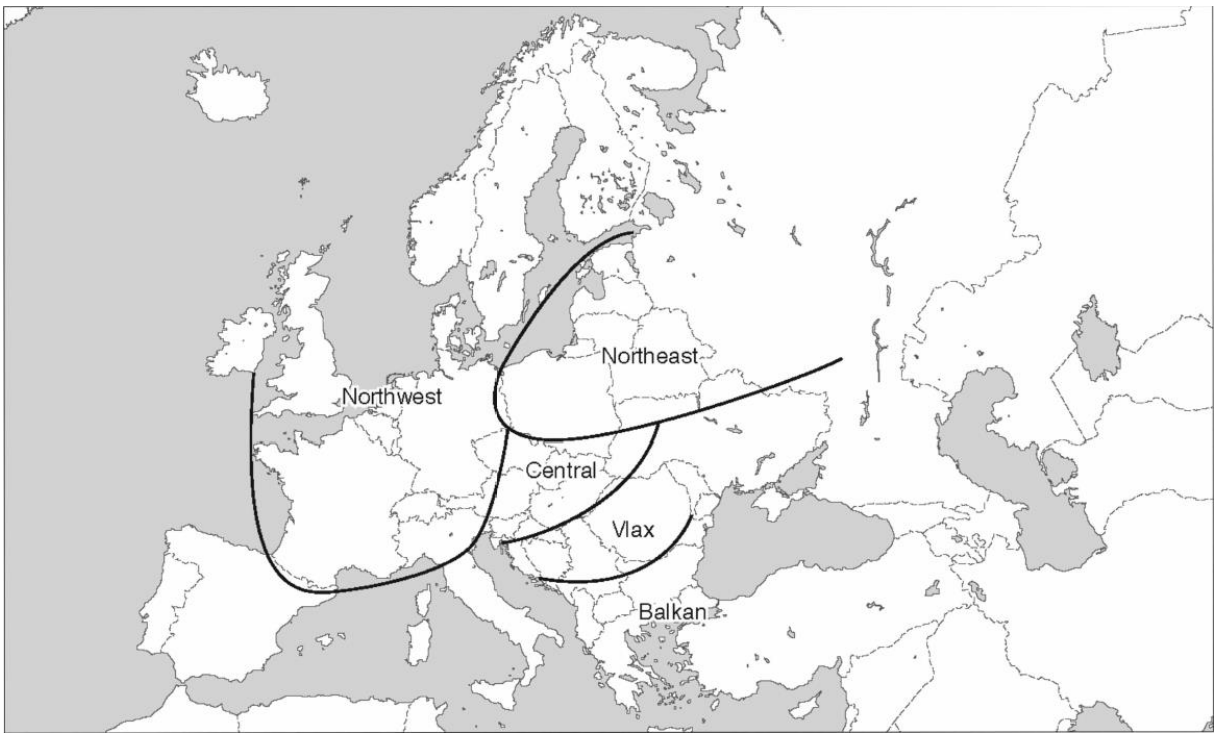


Figure 1 : Classification conventionnelle des dialectes en linguistique romani selon (Matras 2010 : 56)

À cette classification et à la distinction que l'on a vu incontournable entre dialectes dits *vlax* et non-*vlax*, se superpose une ligne de fracture traversant le continent (Figure 2). Il s'agit d'un faisceau très fourni d'isoglosses proposé tout d'abord par Matras (2002 : chap. 9) puis développé par Matras (2005a : 13-17). Il identifie deux épicentres principaux d'innovation en Europe : l'un au nord-ouest de l'Europe (l'Allemagne), l'autre au sud-est de l'Europe, subdivisé en un sous-centre balkanique, et un sous-centre dynamique rayonnant de tous côtés depuis la Transylvanie/Valachie (sous-centre *vlax*). Entre ces deux épicentres principaux passe une ligne baptisée par l'auteur *Great Divide*, ligne de partage. Plus qu'une ligne, il s'agit en réalité d'une zone de transition présentant pour limite nord le sud-est de l'Autriche et la Hongrie, pour limite sud la Slovénie, la Croatie, la Voïvodine et la Roumanie. Cette bande territoriale est étanche : de part et d'autre existent des innovations, mais elles y restent confinées. Elle correspond exactement à la zone de conflit séculaire entre l'Empire ottoman et le reste de l'Europe, et plus précisément à l'aire de transition entre l'Empire ottoman et l'Empire des Habsbourg. La largeur de la zone s'explique par l'importante fluctuation des frontières durant plusieurs siècles. À l'époque où se sont formés les principaux groupes de dialectes du romani, aux XVIe et XVIIe siècle, il s'agissait d'une zone de guerre récurrente, donc d'une frontière politique, économique, sociale, culturelle et religieuse. Cela explique pourquoi cette barrière linguistique semble infranchissable pour les innovations linguistiques : elle l'était pour les locuteurs, qui ont dû cesser d'entretenir des échanges pendant cette période. Étant donné la largeur de la zone, il devait cependant exister des groupes parlant des variétés de transition, présentant donc des caractéristiques à la fois du nord et du sud de la zone. L'existence même de cette ligne de fracture constitue une preuve supplémentaire que les différences dialectales du romani ne remontent pas au romani ancien mais sont nées après la dispersion des Roms en Europe.

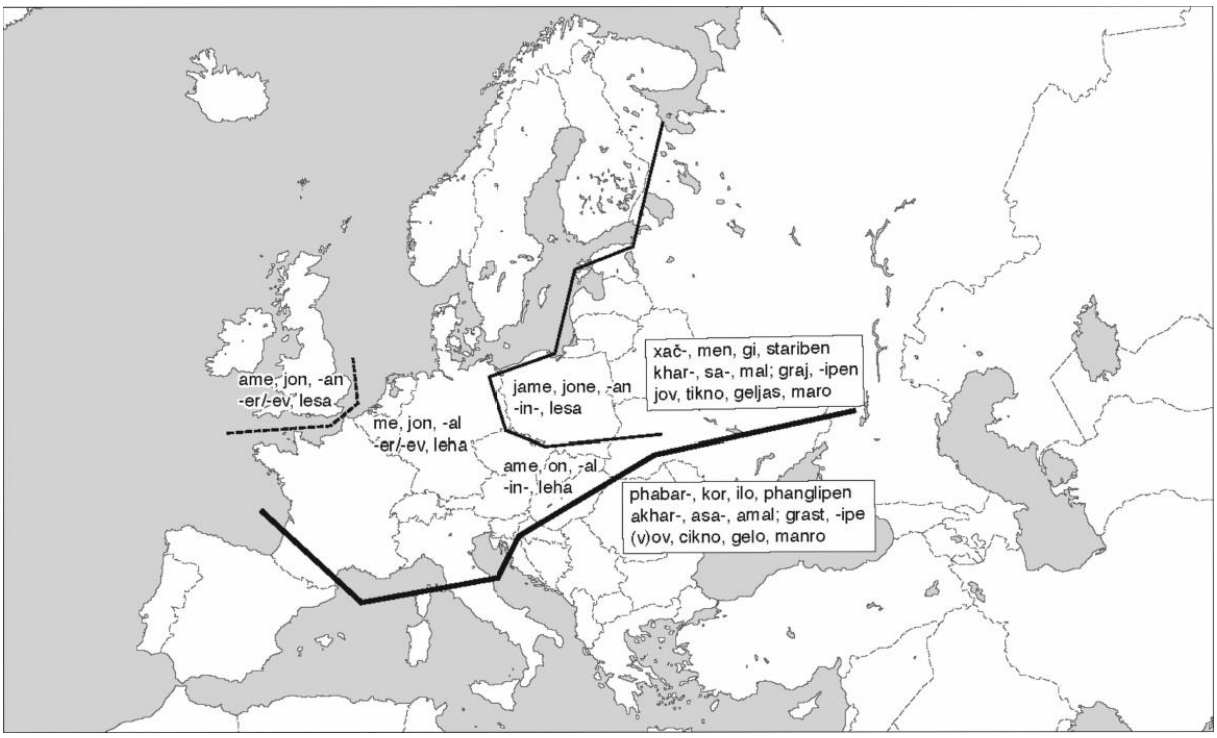


Figure 2 : Les isoglosses majeures et la grande fracture selon (Matras 2010 : 5)

2.2. Le contact de langues

2.2.1. En diachronie

Le romani est une langue indienne balkanisée : « *a hybrid construct of inherited features from Indic, and convergent structures* » (Matras 2000a : 190). Un certain nombre de variétés ont pu être ensuite partiellement débalkanisées. On comprend que le contact de langues soit un élément déterminant dans son étude. L'article de Matras (2000a) est consacré aux changements, parfois extrêmes du point de vue typologique, induits par le contact et qui ont mené à la diversification dialectale. La situation historique et sociolinguistique du romani en fait une langue *en permanence* en contact (Matras 2002 : 191) depuis qu'ils sont Roms, c'est-à-dire depuis que des locuteurs d'une langue indienne ont définitivement quitté l'Inde. Il n'existe aucun locuteur monolingue du romani, à l'exception peut-être des très jeunes enfants : le plurilinguisme est constitutif de la vie des Roms ayant conservé la langue. Les modalités et la pratique du romani peuvent en revanche varier d'un groupe à l'autre. Contrairement à d'autres langues minoritaires, le romani n'a pas fait l'objet, jusque récemment, de politiques volontaristes efficaces pour maintenir et codifier son usage.

Les groupes de Roms ont une histoire sociale, migratoire et linguistique différente : l'histoire des migrations successives et de la vie minoritaire a imprégné les variétés à tel point que l'on peut identifier des « couches » successives d'influences sur le romani. Matras (2002 : 195) propose une stratification diachronique des langues de contact pouvant expliquer les caractéristiques aujourd'hui observables :

- une **L2 ancienne** (*older L2*) « *has had considerable, prolonged impact on the forerunner of a particular dialect. Speakers, especially older speakers, are often aware of this impact, even if the L2 is no longer spoken by members of the community* ».

- une **L2 récente** (*recent L2*) « *is the contact language which is no longer used by the entire community of speakers […] as their everyday language outside the home, but which may still be used by the parent or grandparent generation* ».

- une **L2 actuelle** (*current L2*) « *is the principal contact language used by the community for everyday interaction with the non-Romani majority, and often as a family language alongside Romani* ».

On entendra ici par L2 la langue principale de contact parlée hors de la communauté – mais elle peut également être parlée en famille et peut ne pas être l'unique langue parlée hors de la communauté. On peut la caractériser comme dominante, car c'est celle qui est pratiquée par le plus grand nombre de Non-Roms et parce que c'est celle de l'État, représenté par ses différentes institutions : école, administration, police…

2.2.2. En synchronie

Du fait qu'ils ne sont la population majoritaire d'aucun territoire ni État, les Roms sont tous amenés à côtoyer dans leur vie quotidienne au moins un autre parler local, ne serait-ce que la langue officielle de l'État où ils résident – qui n'est pas forcément le dialecte parlé localement dans le quartier. Cela s'explique par la situation sociale des Roms, toujours dépendants des structures socio-économiques de la société majoritaire. Le romani est circonscrit à la communication interne au groupe, donc restreint à certaines sphères linguistiques. En famille comme en dehors, le choix du romani ou d'une autre langue dépend du contexte de l'interaction : la langue dominante locale est utilisée non seulement avec les Non-Roms, mais également avec les Roms en public, voire en famille. On pourra suivre la répartition tridimensionnelle des usages linguistiques (Denison 1969 et Denison 1971) :

- la langue de l'environnement proche et familier (famille, amis), bref la langue de l'informel dans le domaine privé, est nommée **basilecte**.
- la langue de l'interaction sociale hors du cercle des proches (collègues de travail, connaissances), bref la langue de l'informel dans le domaine public, est nommée **socio-** ou **mésolecte**.
- la langue de l'interaction avec les autorités, l'administration, l'école, la langue des médias, bref la langue du domaine public, est nommée **acrolecte**.

La première est la langue de la connivence, la seconde la langue de la communication standard et la troisième la langue du pouvoir public et idéologique. Ainsi que le fait Deman (2005 : 19) pour le romani de Colombie, nous pouvons schématiser par le Tableau 6 les usages des informateurs interrogés sur le terrain :

Tableau 6 : Schéma du répertoire linguistique des informateurs

Communauté	Basilecte	Mésolecte	Acrolecte
Korçë	ROMANI albanais	ALBANAIS romani	ALBANAIS
Fier	ROMANI	ROMANI (albanais)	ALBANAIS
Llakatund	ROMANI albanais	ALBANAIS romani	ALBANAIS

En majuscules figure la langue principale de l'interaction, en minuscules la langue secondaire, entre parenthèses la langue la moins fréquemment utilisée dans l'interaction. Ces observations personnelles demanderaient bien sûr à être approfondies ou complétées par une une étude sociolinguistique systématique auprès des communautés concernées. Les informateurs de Fier se distinguent des autres car ils vivent dans un quartier strictement rom, où la communication hors du cercle familial peut s'effectuer en romani : c'est pourquoi il figure de façon prédominante en mésolecte. Il figure en revanche de façon mineure chez les Roms des autres villes. Le romani ne peut pas accéder au rang d'acrolecte, car il ne dispose pas (encore) d'une variété standard pan-nationale ni transnationale, et car les Roms n'ont pas l'habitude de l'écrire, ce qui limite ses domaines d'emplois et sa dimension fonctionnelle. La langue minoritaire est donc historiquement et structurellement une langue dominée, dans une relation asymétrique avec la langue dominante, quelle qu'elle soit. Ces facteurs expliquent entre autre que le romani ait peu de prestige auprès des Roms eux-mêmes et soit souvent considéré comme une langue inférieure à la langue dominante (Zatreanu et Halwachs 2003 : 7). On comprend que désir d'ascension sociale aille parfois de pair avec désir d'intégration, voire d'assimilation qui peut aller jusqu'au renoncement à la langue pour certaines familles – ce qui ne signifie pas nécessairement renoncement à la culture et à l'identité rom. C'est selon (Zatreanu et Halwachs 2003 : 7) le cas de Roms de Roumanie, Grèce, Hongrie, Serbie et Roumanie.

Il convient de mentionner ici un phénomène récent, datant de la fin du XXème siècle. Une variété est en train d'émerger chez une partie de l'élite militante européenne, baptisée *khetani ćhib*, langue de rassemblement (litt. langue commune). Elle utilise autant que possible du vocabulaire hérité au lieu d'emprunts aux langues de contact, inévitablement particulières et pouvant oblitérer l'intercompréhension. La morphologie est elle aussi la moins particulière possible : on dira plutôt *som*, je suis, que *siom, sinom, sinjom, isinom, sijum, hom, hum, hium* puisque *som* est le plus petit dénominateur commun entre toutes ces formes. Avec son vocabulaire commun voire avec sa graphie commune, « le rromani commun est d'autant plus apprécié qu'il n'est pas perçu comme lié à un groupe plutôt qu'à un autre » (Courthiade 2014b : 9). L'élite qui l'emploie vise à répandre autant que possible son usage hors des cercles militants. Des expressions issues de variétés particulières peuvent aussi gagner en popularité : les militants choisissent de les adopter lors de leurs échanges internationaux puis de les « emporter » dans leur pays d'origine, leur conférant alors un certain prestige (ainsi *nais tuke* et *palikerav tut*, merci, alors qu'en Albanie on utilise plutôt *ov sasto/sasti* (litt. sois en bonne

santé) ou *faleminderit* (emprunt à l'albanais). En outre, ces locuteurs acquièrent l'habitude de s'exprimer en romani dans un contexte formel, lors des réunions avec un public divers et nombreux, ou encore dans de grandes instances européennes.

2.3. Les variétés des Balkans

Les Balkans ont été la première zone d'Europe continentale par laquelle ont transité les ancêtres des Roms. C'est là que vit encore environ la moitié de la population rom européenne, pour laquelle Asséo (2003) propose une synthèse anthropologique et historique, et plus de la moitié des locuteurs du romani en Europe. Les divisions dialectales les plus importantes de la langue y sont présentes, mais également de nombreux traits conservateurs. Cela en fait donc une zone bien représentative d'un point de vue linguistique. Il faut noter que les frontières des Balkans ne sont pas communément admises : si les limites sud (mer Méditerranée), est (mer Noire) et ouest (mer Adriatique) sont rarement discutées, la limite nord fait débat. La Slovénie, la Croatie, la Voïvodine, la Hongrie, la Roumanie et la Moldavie sont-elles incluses dans les Balkans (Figure 3) ? Cette question n'appelle pas de réponse univoque car elle dépend de l'époque et de la discipline considérées – histoire, anthropologie, sciences politiques… En outre une frontière anthropologique, culturelle ou linguistique ne saurait suivre les frontières des États actuels. Plutôt que d'une frontière nord, il vaut donc mieux rechercher une zone de transition au nord – la meilleure candidate étant celle qui correspond à la zone de guerre entre empires ottoman et des Habsbourg (Figure 2). On considérera, en dialectologie romani, la ligne de fracture de Matras (2005a) comme la frontière nord des Balkans : la Slovénie et une partie du nord de la Croatie, dont l'Istrie en sont exclues, mais le sud de la Hongrie, la Roumanie, la Moldavie et le sud de l'Ukraine y sont incluses – ainsi que la Turquie d'Europe.

Figure 3 : Les Balkans et la ville de Korçë (Albanie)

Le territoire des Balkans ayant été ainsi défini, on trouve dans les Balkans les groupes de dialectes suivants (Figure 1) qui présentent les innovations suivantes (cf. site Romani Project[1]) :

- **balkaniques I** : arli (Macédoine, Kosovo, Serbie, Grèce) et erli (Bulgarie), mečkar (Albanie), sepeči (Turquie et Grèce), ursari (Roumanie), romani de Crimée (Ukraine)…

Ces dialectes présentent les formes *oj, ov, on* pour les pronoms 3SG.F, 3SG.M, 3PL, *m-* comme forme courte du possessif 1SG, des démonstratifs à forme longue *akava/adava*, le suffixe *-in* permettant l'adaptation de verbes étrangers, la négation *na*, la forme *sine* au prétérit 3SG de la copule ainsi qu'un futur formé avec l'auxiliaire invariable *ka*.

- **balkaniques II / *zis*** : drindari, kalajdži, kovački, bugurdži… parlés en Bulgarie et en Macédoine (Boretzky 2000a).

Ces dialectes présentent une fricativisation/affrication et alvéolarisation des occlusives dentales et vélaires (*zis* pour *dis/dives*, jour, *zi* pour *gi/ogi*, âme, cœur, *buci* pour *buti*, travail, *cinel* pour *kinel*, il

[1] http://romani.humanities.manchester.ac.uk/whatis/classification/dialect_classify.shtml

achète…), des démonstratifs à forme courte *kaka*, le suffixe *-iz* d'origine grecque permettant l'adaptation de verbes étrangers.

- ***vlax* du sud** : gurbet, čergar, džambazi, kalburdži (Bulgarie)… parlés en Croatie, Bosnie-Herzégovine, au Monténégro, en Albanie, Macédoine, au Kosovo, en Serbie, Roumanie, Bulgarie, Grèce et Turquie.

Ces dialectes présentent une affrication et palatalisation des occlusives dentales (*dživeh* pour *dis/dives*, jour, *buči* pour *buti*, travail), la forme *dej* pour *daj*, mère, *voj, vov, von* pour les pronoms 3SG.F, 3SG.M, 3PL, *mənr-* comme forme longue du possessif 1SG, des démonstratifs à forme courte *gava*, le suffixe du pluriel *-uri/ura/urja* d'origine roumaine, le suffixe *-isar* permettant l'adaptation de verbes étrangers, la négation *ni* ou *in*, les formes *ses/ha* 3SG et *sem* 1SG au prétérit de la copule, *-em* servant également de désinence prétérit 1SG pour tous les verbes, ainsi qu'un futur formé avec l'auxiliaire invariable *ka*.

- ***vlax* du nord** : kalderaš, lovari, čurari… parlés en Hongrie, Serbie, Roumanie et Moldavie ainsi que partout ailleurs dans le monde en raison des nombreuses migrations de leurs locuteurs (par exemple les Mačvaja parlant une variété de kalderaš, originaires de la plaine de la Mačva en Serbie du nord et qui ont tôt émigré en Amérique du nord).

Ces dialectes présentent une affrication et palatalisation des occlusives dentales (*džes* pour *dis/dives*, jour, *buči* pour buti, travail), une fricativisation (et désaspiration) des affriquées (aspirées) (*žanel* pour *džanel*, il sait, *šavo* pour *čhavo*, garçon, enfant) la forme *dej* pour *daj*, mère, *voj, vov, von* pour les pronoms 3SG.F, 3SG.M, 3PL, *mur-* comme forme courte du possessif 1SG, le suffixe du pluriel *-uri/ura/urja* d'origine roumaine, le suffixe *-i/osar* permettant l'adaptation de verbes étrangers, la négation *či*, la forme *sas* au prétérit 3SG de la copule et la désinence *-em* au prétérit 1SG pour tous les verbes.

- centraux, en réalité seuls les dialectes **méridionaux du groupe central** : romungro, vend, romani du Burgenland… parlés au nord-est de la Slovénie, au sud-est de l'Autriche, au sud de la Slovaquie, en Hongrie, Roumanie et Ukraine.

Ces dialectes présentent une aspiration systématique du /s/ (désinence 2SG *kereha* pour *keresa*, tu feras, *leha* pour *lesa*, avec lui…) voire un amuïssement en finale (*di/dive* pour *dis/dives*, jour), *oj, ov, on* pour les pronoms 3SG.F, 3SG.M, 3PL, des démonstratifs à forme courte *ada*, le suffixe *-in* permettant l'adaptation de verbes étrangers, les formes *sina* 3SG et *sal* 2SG au prétérit de la copule, la désinence *-al* au prétérit 2SG pour tous les verbes.

Les variétés balkaniques du romani présentent des traits linguistiques typiques des langues de cette zone, c'est-à-dire qu'elles participent de l'aire linguistique ou *Sprachbund* balkanique (Sandfeld 1930). On constate pourtant que le romani est peu pris en compte dans les descriptions, pas même dans les synthèses les plus récentes (Feuillet 2012) : c'est le cas de la plupart des « petites » langues non enseignées à l'école et dont aucun État ne revendique les locuteurs. Le trait le plus frappant est la disparition de l'infinitif dans toutes les variétés de romani au monde, ce qui montre que les Roms ont dû, à une époque donnée, stationner ensemble dans les Balkans ou en Anatolie pendant une longue

période[1]. Pour une synthèse sur l'infinitif en romani et ses fonctions, on pourra lire Boretzky (1996). D'autres traits sont le futur analytique formé avec un auxiliaire dérivé du verbe vouloir (*ka* < *kamel*, il veut/aime), l'existence de deux conjonctions de subordination complétives (l'une factive, l'autre non-factive), la réduplication de l'objet par un pronom, l'existence d'un évidentiel passé avec certains verbes, le redoublement de l'objet par un pronom, etc. (Matras et Bakker 1997 : 14).

2.4. Les variétés d'Albanie

Les Roms d'Albanie sont rarement cités dans les études sur le romani, par exemple dans la synthèse effectuée sur les langues parlées par les Roms en Europe centrale et orientale par Marushiakova et Popov (2016). On dispose donc de peu d'informations. Selon Bakker et al. (2000) cité par Bakker (2001 : 299), 95 % des Roms albanais sont locuteurs du romani, ce qui constituait à l'époque le plus haut pourcentage d'Europe. Ils parlent semble-t-il six variétés différentes correspondant aux six groupes ou *fis* (section 1.3, p. 16). Mann (1933a : 4-9) propose une esquisse de phonologie générale du romani albanais, plus exactement des deux – en réalité sûrement trois – variétés qu'il a pu observer sur le terrain. Même s'il est difficile de parler de « romani albanais » commun, notons l'existence du dictionnaire posthume romani-anglais de Mann (1990), édité par Marcel Courthiade à partir du manuscrit de soixante-six pages issu des observations, conversations et des contes collectés en 1929 et 1930 à Tiranë, auprès de Roms de Tiranë, de Shkodër, Korçë, Florina et Thessalonique. Le lexique comprend selon Samara (1990) deux mille entrées avec traduction, explications grammaticales, exemples et expressions idiomatiques. Les variantes « libres ou dialectales » sont proposées en indiquant la région et la variété concernée. La préface indique cependant que sont exclus du dictionnaire « un bon nombre de mots turcs et albanais en usage courant dans ces parlers, afin de mieux mettre en évidence le lexique autochtone rom » (Samara 1990).

2.4.1. Mečkar

Cette variété appartient au groupe dialectal balkanique I. C'est la seule variété albanaise qui soit recensée dans la base de données RMS[2], mais elle n'a pour l'instant pas fait l'objet d'une description. Selon Koinova (2000 : 21-22), cette variété est la plus « consolidée » car c'est celle qui se serait sédentarisée le plus tôt. Elle a donc la plus longue histoire de contact avec l'albanais et le plus grand nombre d'emprunts lexicaux à l'albanais, dont certains très anciens, mais aussi des emprunts au grec. L'influence très importante de l'albanais, qui est aujourd'hui compris de tous les Roms albanais, confère aux locuteurs de cette variété l'impression que leur langue est « gadžéisée » ou en quelque sorte « corrompue ». La transmission de la langue ne semble pas ou semble moins systématique dans beaucoup de familles de ce groupe que dans les autres. Un certain nombre de jeunes adultes mečkars

[1] Certaines variétés ont par la suite recréé un infinitif au centre et à l'ouest de l'Europe, en suivant diverses stratégies morphologiques, notamment en généralisant la forme 3SG du présent (Boretzky 1996). Matras (2000a : 190) parle de *dé-balkanisation*.

[2] http://romani.humanities.manchester.ac.uk/rms/browse

m'ont ainsi exprimé le sentiment d'infériorité qu'ils éprouvent à l'égard de leur variété – sentiment absent chez les Arlis et les Čergars I. Ces observations nécessitent cependant une étude spécifique, approfondie et comprenant un échantillon plus large d'informateurs.

2.4.2. Arli

Cette variété appartient au groupe dialectal balkanique I (Matras 2010). L'arli, présent non seulement en Albanais, mais aussi en Macédoine, au Kosovo, en Serbie et en Grèce, constitue un vaste ensemble de variétés appartenant au même faisceau d'isoglosses, mais présentant une histoire de contact différenciée depuis les dernières décennies. C'est pourquoi nous les traiterons comme des variétés différentes. L'arli de Korçë partage de nombreuses isoglosses avec les trois variétés « arlisques » sud-yougoslaves (Priština, Prizren et Prilep) comparées par Heinschink (1978). Il est également proche des variétés non-*vlax* de Grèce présentées par Matras (2004 : 99-108), notamment de celle de Florina et, en partie, de celle de Parakalamos.

Les emprunts lexicaux de cette variété proviennent majoritairement du turc, et dans une moindre mesure du grec, plutôt que de l'albanais : ils sont donc opaques aux oreilles des locuteurs d'autres groupes, sauf peut-être aux Čergars II. Cela confère à cette variété une certaine aura : celle d'être mieux « conservée » ou moins « corrompue », car moins influencée par la langue dominante locale. Cette variété est selon Koinova (2000 : 21) la seule d'Albanie à avoir emprunté des paradigmes verbaux entiers au turc – phénomène inconnu en Albanie, mais présent en Bulgarie (cf. Draganova 2005), en Grèce (Adamou 2010 : 157-158) et dans les variétés en contact avec le turc, mais aussi le russe et le lituanien (Adamou 2010 : 157). Notons une innovation lexicale relative à l'identité de la communauté : l'existence de deux adjectifs pour désigner la réalité rom, *romano* et *romandikano*). *Romani* s'utilise en opposition binaire à *gadžikani* (pour la langue ou la religion), tandis que *romandikani* renvoie à la réalité rom sans contraste ou antagonisme sous-jacents (ici « la problématique douloureuse »). La communauté rom porte d'ailleurs un nom : *i romandija*.

2.4.3. Čergar I (*vlax*)

Les Čergars I sont originaires de Shkodër au nord du pays, et, auparavant, des zones slavophones de la Yougoslavie. Leur variété appartient au groupe *vlax* du sud, apparentée au gurbet de Serbie et de Bosnie, donc les locuteurs seraient musulmans et partiellement sédentaires (Heinschink 1978 : 11). Stuart Mann publie un corpus très abondant de contes dans cette variété (Mann 1933a, 1933b, 1934a, 1934b, 1941). Il la baptise dialecte du nord, car ses locuteurs viennent de Shkodër et, auparavant, de Serbie. Le gurbet de Bosnie fait l'objet d'un dictionnaire en plusieurs volumes (Uhlik 1941, Uhlik 1942a, Uhlik 1942b, Uhlik 1943a, Uhlik 1943b) et de publications de corpus (Uhlik 1944), celui de Serbie est décrit par Đurić (1983), celui du Kosovo par Boretzky (1986). On peut retrouver les différents gurbets dans l'atlas de Boretzky (2003).

Les emprunts lexicaux de cette variété sont majoritairement faits au slave du sud (serbe ?) et au roman balkanique (roumain ?) plutôt qu'à l'albanais : ils sont donc opaques aux oreilles des locuteurs d'autres groupes. Cela confère à cette variété une certaine aura (et un grand exotisme), par le même phénomène que chez les Arlis.

2.4.4. Čergar II (non-*vlax*)

Selon Marcel Courthiade (c.p.), cette variété appartient au super-dialecte O, c'est-à-dire non-*vlax*. C'est probablement cette variété que transcrit et analyse Stuart Mann dans le conte de son informateur Sali. Mann (1935 : 175) écrit : « *Sali's dialect is more characteristic of the group as a whole. But even in his dialect, although we find undoubted no-Vlachisms, as* me sinóm *I am,* e rakljása *with the girl, as well as* dai *mother (Sc.* dei*),* te les *that thou mayest take (Sc.* te le*),* -esko, esk'i *for* eko, eći, *and the Graeco-Turkish palatals, together with the total abence of Rumanian words, we have also* e rakléa *with the boy,* me phralé *instead of mre phralés,* dák'i *for* dajak'i. *His dialect is full of Turkish, which seems to have had an influence also on its syntax and grammar, for instance his use of the nom. for the acc., and the placing of the verb at the end of the sentence* ». L'expression *e rakléa* pour *e raklesa* pourrait laisser penser qu'il s'agit de mečkar – cependant l'origine du locuteur et la faible présence d'emprunts à l'albanais contredisent cette hypothèse.

C'est probablement aussi la variété du locuteur Neğipi dans Mann (1935, 1937). Celui-ci provient en effet de Macédoine grecque et parle un dialecte non-*vlax* présentant des vlaxismes : « *In this respect the dialect of Neğipi is particularly puzzling. He assures me he is from Salonika, but he says* te le *for* te les, pe *for* pes, pai *for* pani, *and his palatals are quite Graeco-Turkish in character (as in Greek* και, *Turl.* kemal*). Moreover his dialect abounds in Turkish loan-words, as the following tale will show. On the other hand there are no instances of Serb. or Rumanian loans, a sprinkling of Greek and Bulgarian words being used in their stead.* » (Mann 1935 : 175).

2.4.5. Bamil et kurtofi

Selon Marcel Courthiade (c.p.), le bamil appartient au super-dialecte O, en d'autres termes il s'agit d'une variété non-*vlax*. Selon Koinova (2000 : 21) qui s'appuie sur l'opinion de Marcel Courthiade, le kurtofi appartient au groupe balkanique – il s'agit donc là aussi d'une variété non-*vlax*. Les plus anciens membres de ce groupe parleraient plus volontiers grec que romani.

2.4.6. Les néo-locuteurs

Le concept de *new speaker*, néo-locuteur, a été forgé par Bernadette O'Rourke (voir par exemple O'Rourke et Ramallo 2013) pour désigner les locuteurs de langues régionales minoritaires à qui la langue n'a pas été transmise par la famille ou la communauté, mais par l'école, suite à la mise en place

de programmes volontaristes de revitalisation. C'est le cas pour le gaélique d'Irlande ou le galicien. Elle parle également de « *'non-native' newcomers to the language* » (O'Rourke et Ramallo 2011). Le contexte sociologique est bien sûr tout à fait différent : le romani d'Albanie ne semble pas en voie d'extinction, il ne fait pas l'objet de programmes de revitalisation ni de programmes tout court, scolaires ou autre, de la part de l'État. Mais l'on peut considérer que le processus à l'œuvre ici est en partie comparable, car la transmission n'a pas été effectuée dans le cadre familial, ce qui amène ces personnes à apprendre ou à approfondir la langue plus tard, dans un cadre scolaire, dans une démarche consciente et active – et personnelle, contrairement aux écoliers d'Irlande ou de Galicie.

2.4.7. Contact de langues

À partir de l'histoire migratoire des groupes de Roms ainsi que des emprunts lexicaux attestés dans leurs variétés, on peut présumer l'histoire du contact telle que présentée dans le Tableau 7. Le contact avec les langues entre parenthèses est considéré comme occasionnel et, en ce qui concerne la L2 actuelle, dû aux migrations et/ou aux médias (radio, télévision, Internet). Toutes les variétés sont actuellement en contact avec l'albanais et ont été en contact avec le grec par le passé. Les Roms parlent ou parlaient comme L2 récente les autres langues parlées dans la région où ils vivent : grec, turc, slave du sud (serbe, macédonien, bulgare). On note que les Bamils constituent le groupe le plus récemment arrivé, et que les Arlis et les Čergars II partagent la même histoire migratoire, donc la même histoire de contact récente. Les Čergars I ne se distinguent que par les L2 anciennes, qui comprennent le bulgare, si l'on considère leur région d'origine, et en supposant qu'il s'agit de la variété de Neğipi décrite par Stuart Mann – auquel cas il faudrait d'ailleurs ajouter le romani *vlax* comme variété de contact ancienne.

Tableau 7 : Hypothèse d'histoire du contact de langues des variétés de romani d'Albanie

	Mečkar	**Arli**	**Čergar I**	**Čergar II**	**Bamil**	**Kurtofi**
L2 anciennes	albanais grec (slave du sud ?)	turc grec	langue romane (slave du sud)	turc grec (slave du sud)	grec (turc)	?
L2 récentes	albanais (grec)	grec albanais (turc)	slave du sud albanais (grec)	grec albanais (turc)	grec	?
L2 actuelles	albanais (grec) (italien)					

3. Méthodologie générale

Il existe trois grandes méthodes permettant d'étudier un phénomène linguistique en typologie : l'introspection et les jugements de grammaticalité, l'étude sur corpus (dans une perspective qualitative et/ou quantitative), la réalisation d'une expérience psycholinguistique (dans une perspective en général quantitative). Toutes trois donnent un éclairage différent du phénomène linguistique. Partant du même constat que Samvelian (2017), nous avons choisi d'employer une méthodologie empirique pour notre étude, c'est-à-dire de collecter et d'analyser un corpus de données produites par des locuteurs natifs en contexte écologique, ainsi que de mettre en place une « expérience » linguistique.

> « L'étude de la syntaxe a été longtemps dominée par une approche opposant "la grammaticalité" à "l'agrammaticalité", l'objectif étant alors de modéliser la « compétence » des locuteurs afin de rendre compte des phrases bien formées en excluant celles mal formées. Pour atteindre cet objectif, la syntaxe formelle, en particulier le modèle génératif chomskyen, a eu essentiellement recours à l'introspection, celle du linguiste travaillant sur une langue où « les locuteurs natifs » interrogés par ce dernier quant à leurs jugements de grammaticalité sur des énoncés de leur langue. Cette méthodologie, qui peut à la rigueur convenir pour la description des phénomènes manifestant un contraste catégorique clair entre les données grammaticales et agrammaticales (ex. en français le déterminant défini précède toujours le nom qu'il détermine), s'avère problématique dès que des constructions plus complexes pour lesquelles les jugements de grammaticalité peuvent être incertains ou variables d'un locuteur à l'autre, d'une situation à l'autre, etc. Un autre problème est posé par des phénomènes qui impliquent une variation ou la possibilité pour les locuteurs de choisir entre deux ou plusieurs possibilités. […][1] Pour chacun de ces phénomènes, les deux possibilités sont équivalentes du point de vue de leur grammaticalité. Toutefois, lorsque l'on se penche sur la production (performance) des locuteurs, l'on constate qu'il existe des préférences en fonction de paramètres précis. Ainsi, comme il a été montré dans plusieurs études des paramètres tels que la longueur, la complexité syntaxique, le statut informationnel (topic vs focus), le statut référentiel (défini vs indéfini) interviennent dans toutes les langues pour déterminer le choix d'un ordonnancement par rapport à un autre. Ces préférences, qui font également partie de la compétence linguistique des locuteurs, peuvent faire l'objet de modélisation avec les méthodes empiriques (corpus, expériences…). » (Samvelian 2017)

3.1. Questionnaire et élicitation ?

Pour étudier les syntagmes polydéfinis, on pourrait partir de l'intuition des locuteurs, en procédant par élicitation, par traduction de phrases ou par questionnaire de jugements de grammaticalité. Cette méthode, basée sur l'introspection des locuteurs et reposant sur la distinction entre compétence et performance, constitue la source de la grande majorité des travaux en syntaxe. Elle consiste à demander « l'avis » des locuteurs quant à l'appartenance de structures ou de phrases au système de la langue, en faisant appel à leur intuition. Mais « il est important de préciser le statut du jugement de grammaticalité : les locuteurs natifs ne fournissent pas des jugements de grammaticalité, dans la

[1] Telles que la réalisation facultative de *that* dans certaines subordonnées complétives ou relatives en anglais, l'alternance dative en anglais ou encore l'ordre libre des constituants dans certaines langues ou pour certains syntagmes.

mesure où la notion de grammaticalité est une construction théorique. En effet, les intuitions des locuteurs ne sont que des jugements d'acceptabilité et c'est au linguiste de décider si ces jugements sont le reflet d'une propriété de la compétence ou de phénomènes liés à la performance » (Thuilier 2012 : 26). La grammaticalité est donc une construction théorique, qui fait appel à des règles morphologiques, syntaxiques, lexicales, sémantiques : il est donc nécessaire de la distinguer des intuitions des locuteurs. Il incombe ensuite aux linguistes, professionnels spécialisés de la langue, d'interpréter les jugements d'acceptabilité (Schütze 1996 : 26). Le processus, résumé en Figure 4, est le suivant : les locuteurs fournissent des intuitions qu'ils convertissent en jugements d'acceptabilité. Les linguistes convertissent ces derniers en jugement de grammaticalité, sur la base d'une décision théorique de nature hypothético-déductive, partant de l'observation de propriétés afin d'émettre une hypothèse générale. « Ces propriétés sont notamment les contrastes entre phrases jugées acceptables et inacceptables, que le chercheur interprète comme des contrastes de grammaticalité » (Thuilier 2012 : 27). De là, le linguiste formule des hypothèses explicatives pour ces contrastes, et pour rendre compte des jugements de grammaticalité : il fonde ainsi sa description du système de la langue.

intuition > jugements d'acceptabilité > décision théorique hypothético-déductive > jugement de grammaticalité > hypothèse(s) explicative(s) permettant de saisir ces contrastes > description du système de la langue

Figure 4 : Processus permettant de décrire une langue sur la base de l'intuition et des jugements d'acceptabilité

Cette méthode de recherche est complémentaire de l'étude de corpus de données spontanées et de l'expérience psycho-linguistique produisant des données semi-spontanées. Si l'analyse des jugements de grammaticalité permet de définir les contraintes catégoriques qui structurent le système d'une langue, ce sont les deux autres qui permettent « l'étude et la mise à jour de contraintes préférentielles qui agissent dans les zones de liberté offertes par la structure syntaxique fondamentale de la langue » (Thuilier 2012 : 33-34) En revanche, l'utilisation de la seule intuition et des jugements pose deux problèmes principaux :

- le manque de contrôle des conditions dans lesquelles le jugement est produit, soit l'absence de véritable protocole expérimental « mis en place de façon à contrôler les facteurs extra-linguistiques pouvant avoir une influence sur les jugements »[1] (Thuilier 2012 : 27),
- les données elles-mêmes soumises au jugement, qui sont des « phrases construites et décontextualisées. Cela peut conduire au rejet de certaines structures qui pourraient pourtant être produites dans des contextes plus élaborés »[2] (Thuilier 2012 : 27).

De plus, certaines formulations, qui s'écartent des formulations communes et paraissent étranges ou « accidentelles », sont parfois produites en discours et méritent donc d'être prises en compte. Une

[1] (Schütze 1996 : 52-53) identifie trois types de problèmes relatifs à cette méthode : les jugements de grammaticalité ne sont pas systématiquement reportés et la notation n'est pas unifiée ; certains jugements sont parfois retenus ou écartés ; certains jugements demandent aux locuteurs des capacités spécifiques, sans preuve qu'ils soient réellement capables d'effectuer les distinctions demandées, et sans aucun contrôle du processus d'obtention des jugements.

[2] Plusieurs travaux ont montré comment les jugements de grammaticalité rencontrés dans les travaux de syntaxe « sous-estiment l'espace de possibilité grammaticale » selon Bresnan (2007 : 1), citée par (Thuilier 2012 : 27).

analyse fondée sur le seul jugement de grammaticalité oblitérerait au contraire leur existence. Dans le cas d'une structure optionnelle, qui ne fait donc pas l'objet de contraintes catégoriques mais apparaît pour des raisons pragmatiques, une démarche par jugements d'acceptabilité, élicitation ou traduction semble moins adaptée à notre sujet d'étude. À moins de créer une situation d'énonciation artificielle en « jouant » chaque scène à l'oral afin de créer un contexte pertinent, il manque en effet une situation d'énonciation authentique. Partant, il est difficile d'interpréter les résultats. Les tâches par élicitation sont selon nous appropriées pour dresser des paradigmes morphologiques ou syntaxiques, mais échouent à prédire ce qui est, dans l'absolu, possible ou non dans la langue.

Toutes ces objections tendent à remettre en cause la validité des données issues de ces méthodes traditionnelles. Nous pensons également que la seule introspection – même celle du chercheur – ne suffit pas pour décrire adéquatement la langue ni pour proposer un modèle théorique à l'épreuve de la réalité, car l'analyse de la performance est marquée par la subjectivité du chercheur. De plus, n'étant pas locutrice native du romani, je ne peux me fier tout à fait à mon intuition ni à ma compétence, nécessairement parcellaire. Il existe en outre d'autres obstacles concernant le romani.

- La très grande majorité des Roms, même alphabétisés, n'a aucune expérience du romani écrit. Ils ne l'écrivent pas eux-mêmes et ne l'ont jamais lu. Le fait de ne jamais avoir vu une langue sous forme écrite la rend quasiment étrangère au premier abord.
- Cette langue n'est pas enseignée à l'école en Albanie. Cela implique que les locuteurs n'ont jamais assisté à des cours de grammaire, donc jamais été initiés à une approche méta-linguistique systématique de leur propre langue. Un questionnaire, même oral, faisant appel à un jugement métalinguistique est donc une expérience nouvelle. Un travail d'élicitation ou de traduction systématique est une expérience difficile sur le plan cognitif – à moins que le locuteur ne soit un militant qui donne des cours d'alphabétisation, voire des cours de langue romani dans un cadre associatif.
- Ce travail est abstrait, hors contexte, ce qui est une difficulté très importante pour la plupart des locuteurs, peu alphabétisée. Il n'est possible qu'auprès de locuteurs instruits, par ailleurs patients et bienveillants.

Il va de soi qu'une telle démarche permet d'éclairer non pas tant les pratiques réelles des locuteurs que leurs pratiques déclarées et leur conscience linguistique. J'ai souhaité toutefois tenter l'élicitation afin de contrôler l'existence de structures, comme DEF A DEF N, qui ne sont pas ou peu apparues dans le reste des données, par contingence ou par impossibilité grammaticale. Cela me permettait également de mesurer la variation dialectale, en testant des structures que je savais exister dans d'autres variétés comme celles de Roumanie. N'étant pas locutrice native de la langue, j'ai établi une liste de phrases à partir d'exemples cités dans les publications scientifiques ainsi que d'exemples tirés du recueil d'interview en romani roumain de Cioabă (2006a) en traduisant les emprunts par des équivalents albanais. À partir de cette série d'expressions, j'ai élaboré un questionnaire de grammaticalité visant à tester la place de l'adjectif, son type sémantique, le nombre d'adjectifs, le type et le nombre de déterminants. Le fait d'expliquer le but de l'élicitation en romani et de ne pas demander aux locuteurs de traduction réduisait les effets du contact de langues et les « calques » de la langue dominante, notamment lorsque je demandais de produire des exemples. Ma collaboratrice

la plus proche, une jeune femme scolarisée, F16AČ, a été la première à se livrer à cette tâche. Je lui ai demandé de m'indiquer si elle considérait les phrases comme acceptables et lui ai demandé de les répéter à voix haute, afin d'obtenir une confirmation orale et personnelle (Sakel et Everett 2012). Je lui ai proposé si besoin, de corriger les phrases et d'inventer des exemples qu'elle considérerait plus pertinents. Cette tâche a été réalisée dans une pièce isolée, au moyen de feuilles papier et d'un fichier informatique de traitement de texte, où la locutrice et moi écrivions, modifions et biffions les phrases au fur et à mesure du questionnaire et de la discussion. À la demande de ma collaboratrice, les phrases étaient écrites en romani orthographié selon les conventions de la langue albanaise. L'ensemble de la conversation a par ailleurs été enregistré.

Les résultats furent laborieux et peu convaincants. La locutrice a rejeté des formes en principe acceptables car entendues par ailleurs sur le terrain. Ce rejet peut avoir diverses causes : manque de concentration, d'habitude d'effectuer des raisonnements métalinguistiques (surtout en romani sur le romani), lassitude, ennui… Elle a au contraire accepté des formes « en principe » inacceptables, ce qui ne s'explique pas que par la fausseté de mes prédictions et le manque de concentration et d'habitude des raisonnements métalinguistiques. Il est attesté que le désir de faire plaisir au chercheur peut fausser les résultats (Spike 2014). J'ai moi-même constaté dans la vie courante (monolingue en romani) que le seul fait qu'une expression soit compréhensible, que le contenu informatif soit décodé par l'interlocuteur peut amerner celui-ci à considérer l'expression comme correcte donc « acceptable ». C'est notamment le cas pour une langue non normée comme le romani, non enseignée et soumise à une variation dialectale très importante. Les locuteurs, notamment s'ils fréquentent des Roms d'autres groupes albanais voire d'autres pays, sont habitués à entendre une grande diversité de prononciation et de différences morphologiques, syntaxiques et lexicales. Les notions de « norme » et de « grammaticalité » ont donc une autre dimension et prennent un autre sens lorsque l'on travaille sur le romani (cf. entre autres Matras 2005b et Matras 2015). Les études par élicitation auprès des locuteurs natifs concernent – presque – toutes de « grandes » langues, parlées sur un territoire compact, par un nombre important de locuteurs, le plus souvent scolarisés, des langues présentant une variété standard et/ou des institutions reconnues et/ou un enseignement scolaire. Ce n'est pas le cas du romani, ce qui constitue un obstacle à l'utilisation des tests de grammaticalité.[1] Ces tâches mobilisent une connaissance déclarative, consciente voire normative de la langue, alors que le discours spontané mobilise une connaissance procédurale implicite et moins consciente de la langue. Il ne m'a donc pas semblé opportun et pertinent d'utiliser l'intuition et le jugement d'acceptabilité. Le recours aux grammaires et aux descriptions (dialectales) est quant à lui non seulement problématique, mais aussi impossible : il n'en existe pas pour l'Albanie, ni assez pour couvrir l'ensemble des variétés (même balkaniques).

J'ai donc renoncé à poursuivre dans cette voie. De plus, je ne pouvais effectuer une telle tâche avec beaucoup de locuteurs car il était très difficile de trouver une personne disponible pour une tâche longue, abstraite et aride, requérant beaucoup de concentration et de recul. Cela limitait le processus de comparaison. Je préférais par ailleurs ne pas faire connaître la structure grammaticale qui faisait

[1] Il existe bien sûr des recherches menées par élicitation et traduction sur des langues minoritaires ayant le même profil que le romani (cf. par exemple l'immense travail de Durand (2016) qui porte sur les langues amérindiennes arawak et a fait l'objet de plusieurs séjours de recherche en Amazonie : l'auteur a procédé par élicitation et traduction), mais elles font alors l'objet d'une rémunération, ce qui n'a pas été le cas de la mienne. De plus, un tel rapport « salarié » avec les locuteurs est loin d'assurer la validité des résultats.

l'objet de mon étude, afin de ne pas influencer la façon de s'exprimer des personnes qui m'entouraient. C'est d'ailleurs pourquoi je n'ai fait passer ce questionnaire qu'à la fin de mon deuxième séjour albanais.

3.2. Linguistique de corpus

J'ai choisi de partir d'un ensemble d'hypothèses, de prédictions que la réalité viendrait ensuite infirmer ou confirmer, en travaillant à partir d'un corpus. Comme l'écrit Thuilier (2012 : 31), « l'utilisation de corpus annotés accompagnés de méthodes d'analyse de données adéquates permet d'apporter des observations nouvelles et complémentaires pour décrire et comprendre un ensemble de phénomènes syntaxiques ». Il existe deux manières de recueillir un corpus de données. La première méthode consiste à faire se déplacer les locuteurs dans un bureau ou un laboratoire où le chercheur contrôle les conditions d'enregistrement, d'énonciation, le profil et le nombre des locuteurs, la durée de la conversation, son contenu… Le déplacement et la collaboration des locuteurs sont alors bien souvent rémunérés. La seconde méthode consiste pour le chercheur à aller trouver les locuteurs dans leur milieu, ce qui n'est pas sans obstacles ni dangers. Le corpus est alors restreint au bon vouloir des locuteurs, à leur disponibilité, à leur désir de collaborer ou non. Les locuteurs peuvent ne pas lui dire la vérité ou la déguiser. Il ne peut en outre jamais rencontrer et enregistrer toute la communauté, mais seulement les personnes que peuvent ou veulent lui faire rencontrer ses plus proches collaborateurs, ou ceux avec lesquels il a le plus d'affinités. C'est pourquoi le travail en immersion, même s'il se donne pour but d'être plus réaliste ou « écologique », reste empreint de subjectivité : il décrit le parler du réseau de personnes que fréquente ou qu'apprécie le chercheur. Entre ces deux pôles, dont l'un recherche plutôt le confort, la rapidité et l'efficacité et l'autre plutôt l'authenticité, existent de nombreuses situations intermédiaires. On peut par exemple effectuer des recherches de terrain sans immersion, lorsque l'on ne vit pas dans la communauté mais que l'on rencontre occasionnellement ses informateurs dans un lieu neutre (ni son laboratoire, ni le lieu de vie de la communauté) : café, salle associative… Enfin, certains chercheurs effectuent des cyber-recherches depuis leur bureau ou leur domicile. Nous allons donc dans le même sens que Fagard (2015 : 13-14) :

> « L'étude sur corpus permet ainsi de dépasser la question assez problématique de l'acceptabilité. […] Par ailleurs, l'utilisation de tests d'acceptabilité dans la littérature me semble souvent peu probante, du moins lorsqu'elle est formulée en ces termes : bien souvent, il est assez facile de trouver un contexte où la phrase réputée inacceptable devient tout à fait naturelle, sans pour autant nécessairement remettre en question l'intérêt ou la solidité des résultats. Cela tient au fait que la question la plus intéressante n'est pas celle de la possibilité ou non d'une construction donnée. Le plus intéressant, me semble-t-il, est plutôt de savoir si une construction ou une séquence est naturelle, usuelle, en d'autres termes d'évaluer ses probabilités d'apparition. » (Fagard 2015 : 13-14)

L'un des buts de notre recherche est en effet de comprendre quand et pourquoi les locuteurs emploient la structure rare qui fait l'objet de ce travail. On pourrait cependant émettre des réserves quant au fait de se fier à un seul corpus :

- **Manque de représentativité de l'échantillon de personnes enregistrées** : il n'existe pas d'étude sociologique sur la population rom albanaise. Le corpus ne saurait donc prétendre

représenter les locuteurs en respectant les proportions de l'âge, du genre, du niveau social, etc. Cela rejoint la question de Thuilier (2012 : 31) : « comment passer d'observations effectuées sur une communauté de locuteurs à la connaissance détenue par le locuteur individuel ? Autrement dit, en quoi les fréquences observées en corpus sont en rapport avec la production d'un locuteur ? Pour tenter de répondre à cette question, nous émettons la conjecture selon laquelle l'usage façonne, au moins en partie, la connaissance des locuteurs. »

- **Manque de représentativité du corpus** : il n'existe pas de base de données suffisamment conséquente en romani, qui recense à la fois la langue orale et la langue écrite (lorsqu'elle existe), couvrant un grand nombre de variétés dialectales et un grand nombre de registres (littéraire, spontané, semi-spontané…). Il en existe plusieurs pour des langues de grande diffusion et bien décrites comme l'anglais, le français, l'allemand… mais non pour le romani. Cependant, il n'existe pas de corpus idéal, présentant tous les genres écrits et oraux possibles, tous les contextes de conversation possibles, formels et informels. Aucun corpus ne saurait couvrir la totalité de la langue existante et énoncée, ni de la langue possible, car un corpus n'est jamais qu'un infime échantillon de l'ensemble des phrases produites dans une langue donnée (Fagard 2015 : 15, Thuilier 2012 : 45-46).[1] Cela rejoint la question de la généralisation (Thuilier 2012 : 44), à savoir « comment passer des observations sur corpus à des propriétés du système de la langue ? ».

Il n'existe pas de langue romani standard dont les données langagières seraient aisément répertoriables dans une base de données. C'est pourquoi la fondation de corpus en romani doit passer, et commencer, par la dialectologie. Le seul corpus en ligne existant à ce jour est la *Romani Morpho-Syntax Database* (Matras, Elšík et al. 2001, Matras, White et Elšík 2009). Cette œuvre collective consiste à recenser et à étudier le plus grand nombre possible de dialectes dans le monde. On y trouve une liste de plus de 1062 mots et phrases traduits en romani depuis la langue de contact locale par des locuteurs natifs, mais aussi des entretiens libres et des exemples issus de certaines publications scientifiques. Les données se présentent sous format audio et texte (après transcription et traduction). De nombreuses isoglosses, notamment phonologiques, sont visualisables sur des cartes dynamiques ; un moteur de recherche interne permet de faire une recherche inter-dialectale sur une partie du discours ou un point de grammaire particulier. Un seul dialecte albanais a été à ce jour recensé sur la base de données : le mečkar, consultable en sélectionnant AL-001 – Mečkari sur cette page[2]. La glose interlinéaire des exemples est disponible sur cette autre page[3]. Ce corpus pourra servir de point de comparaison avec le nôtre (cf. section 1.3.2, p. 411) – mais ne peut pas nous servir de base, car nous préférons travailler sur des données orales spontanées (cf. section 3.4 p. 53).

[1] Sur la question de la représentativité d'un corpus linguistique, de l'échantillonnage aléatoire stratifié et de la recherche de l'absence de biais, cf. Evert (2006) et Thuilier (2012 : section 2.1.1). Il est impossible de connaître le nombre, la diversité et la taille des textes constituant un corpus représentatif.

[2] https://romani.humanities.manchester.ac.uk/rms/browse/phrases/phraselist

[3] https://romani.humanities.manchester.ac.uk//rms/browse/transcript/transcripts

3.3. Linguistique de terrain

3.3.1. La question de la maîtrise de la langue

Souhaitant étudier la langue dans son contexte réel et collecter des données de « première main », j'ai opté pour la seconde méthode, le travail de terrain *in situ*, consistant à me rendre en Albanie et à vivre au sein de la communauté. N'étant pas locutrice native du romani, je me suis d'abord initiée à la langue, à l'histoire et à la culture des Roms en passant le diplôme de licence de *rromani* à l'Inalco en 2010-2013. Il était à l'époque constitué des enseignements de langue de Marcel Courthiade, Saimir Mile et Orhan Galjuš ainsi que de divers enseignements d'histoire et de culture indienne et balkanique. Le travail, même sur un point grammatical particulier, requiert selon moi un certain niveau de maîtrise de la langue étudiée. Même s'il n'y a pas de lien direct entre le niveau de maîtrise d'une langue et la qualité des analyses menées, on peut penser que la finesse de celles-ci est conditionnée par la connaissance de la langue (Fagard 2015 : 29). Le fait de ne pas être locutrice native de la langue constitue, de ce point de vue, un grand inconvénient – compensable par l'étude préalable de celle-ci. Mais elle permet d'éviter la fausse aisance et donne du recul sur son fonctionnement grammatical.

Marcel Courthiade[1] m'a recommandée et mise en contact avec mon hôte de Korçë, Arben Kosturi. Comme l'anthropologue Enkelejda Sula Raxhimi, je pense que « travailler dans les quartiers roms et chez les Roms n'est pas évident, surtout pour une non-Rom comme moi » (Sula Raxhimi 2015 : 58). C'est pourquoi il est nécessaire de faciliter les contacts en se faisant introduire dans une famille ou une communauté par l'un de ses membres. La famille d'Arben Kosturi parle les variétés arli (de son côté) et čergar I (du côté de son épouse), l'albanais (des deux côtés) et, dans une moindre mesure, le grec (des deux côtés). J'ai réalisé deux séjours en immersion monolingue à Korçë (cf. Figure 3) : trois semaines en 2013 puis quatre semaines en 2014.

3.3.2. Le déroulement des séjours

J'ai abordé la recherche sur le terrain avec une approche ethnographique. Il s'agissait d'une recherche en immersion, avec une première phase d'imprégnation et d'observation participante qui m'a permis de m'adapter à la variété et à la culture locale. La vie quotidienne monolingue en romani m'a amenée à une utilisation bidimensionnelle du romani, utilisé à la fois comme basilecte (cf. section 2.2.2, p. 35) dans la famille qui m'accueillait, comme mésolecte dans les rapports avec les personnes interviewées ou lors des réunions de travail nationales ou internationales. J'ai opté pour des enquêtes monolingues afin d'inciter les locuteurs à pratiquer au maximum cette langue en ma présence et de minimiser, si possible, les interférences entre le romani et la langue localement dominante dans leurs interactions avec moi. Lors de ces séjours, j'ai tenté de m'inclure dans la communauté afin d'essayer d'atténuer mon statut d'étrangère : en essayant de créer un climat de confiance et d'intimité favorisant des interactions langagières qui soient le moins possible perturbées par ma présence. C'est aussi

[1] Maître de conférences en langue et civilisation rromani à l'Inalco, Paris.

pourquoi j'ai volontiers partagé les tâches à la maison comme au travail, ainsi que le mode et le rythme de vie des familles qui m'accueillaient.

La famille de Korçë est la plus respectée de la communauté locale. Elle est célèbre parmi les familles roms en Albanie, tant sur le plan du prestige que de l'éducation. Mon hôte m'a accueillie (Figure 5) et présentée à sa famille avec bonté et générosité et m'a présentée à de nombreux locuteurs de sa communauté à Korçë, mais également de Tiranë, Fier, Berat, Kuçavë, Llakatund, Pogradec, Bilisht et Durrës grâce à son réseau étendu de relations. Il m'a permis d'assister, en juillet 2013 et partiellement en juillet 2014, au congrès (*akademi*) annuel des minorités albanaises à Voskopojë : j'ai pu y rencontrer des Égyptiens, Macédoniens, Serbes, Vlah (en albanais ou romani *Çobane*) et LGBT du pays. À l'été 2013, j'ai pu participer à deux congrès sur les droits et la langue des Roms et Égyptiens à Tiranë. Nous nous sommes rendus, en 2014, à une réunion militante à Ohrid (ancienne République yougoslave de Macédoine) pour le projet *Days of Roma Culture* du programme *Social Inclusion of the Roma – Challenge for Regional Cooperation and Networking in the Cross Border Region Macedonia-Albania*. Lors de ce séjour, j'ai eu l'occasion d'enregistrer des personnes parlant arli, mečkar et čergar I, mais aussi deux Roms macédoniens locuteurs des variétés džambaz et kovački et un Rom kosovar vivant en Allemagne, locuteur de la variété xoraxane. Enfin, j'ai eu la chance de participer, en 2013, à la grande fête de mariage du neveu de mon hôte ainsi qu'aux préparatifs de celui-ci et, en 2014, à l'anniversaire de mon hôte, qui est chaque année l'occasion d'une grande réunion de famille.

Figure 5 : Vue des alentours de la maison de mes hôtes à Korçë

En contrepartie, j'ai participé au travail de journaliste, aux réunions et aux reportages télévision de mon hôte : j'ai présenté et commenté des extraits vidéo à l'écran, et répondu à une interview. Il m'a permis de l'accompagner dans diverses villes pour interviewer des Roms et documenter leur situation sociale dramatique. Je lui ai apporté une aide linguistique, par exemple pour écrire en anglais ou pour traduire des documents du romani à l'anglais. En 2013, je suis passée quatre fois à la télévision nationale TV Klan avec sa fille, journaliste dans une émission quotidienne, dans le but de faire la promotion de la langue romani. Nous avons également ébauché un projet de documentaire sur la culture des Roms albanais et sur leurs problèmes socio-économiques, ainsi qu'un projet de description dialectologique du romani d'Albanie, là aussi dans une perspective de vulgarisation.

La recherche initiale envisageait plusieurs sites de recherche dans plusieurs pays, afin de comparer plusieurs variétés balkaniques entre elles. J'ai donc effectué une petite mission en Serbie, à Zemun, pendant dix jours en 2013. La famille parle une variété gurbet (sud-serbe), la variété arli (kosovare), le serbo-croate et, selon les divers locuteurs, le russe, l'anglais et le suédois. J'ai enfin effectué un séjour en Grèce (Thessalonique, Larissa et Veria) pendant deux semaines en août 2014, qui m'a permis d'effectuer des entretiens auprès de nombreux locuteurs Roms grecs et immigrés bulgares et albanais. Le présent travail comporte seulement les résultats du travail en Albanie.

3.3.3. Le rôle de chercheuse sur le terrain

Le rôle du chercheur sur le terrain est complexe et polymorphe. Sa présence modifie l'écologie sociale (question, par exemple, du dédommagement pour l'hébergement et de la rémunération des collaborateurs) et linguistique. Des membres de la communauté peuvent s'efforcer de parler, plus qu'à l'ordinaire, la langue qui intéresse le chercheur, soit pour l'aider dans sa recherche et lui rendre service, soit parce que le chercheur n'a pas d'autre langue en commun avec eux. Comme je parle très peu l'albanais, certains Roms étaient contraints de s'exprimer avec moi en romani dans des contextes où ils se seraient spontanément exprimés dans la langue majoritaire (contextes de mésolecte voire d'acrolecte). Paradoxalement, j'ai pu donc être à l'origine d'une modification (intensification quantitative voire élargissement qualitatif) de l'usage de la langue.

Sur le terrain, j'ai souvent eu l'impression d'être « hors genre ». Les activités que j'effectuais et le rôle que je revêtais me classaient d'ailleurs en fait plutôt du côté masculin : mon travail m'amenait à sortir de la maison et à voyager à travers la ville et le pays beaucoup plus que la plupart des femmes rom (mais pas autant que les hommes) – ce qui me fut confirmé par le ressenti d'un certain nombre de membres de la communauté. J'ai pu donc participer à des activités considérées comme masculines (travail intellectuel, informatique, sorties le soir, certains préparatifs du mariage) autant qu'à des activités considérées comme féminines (travail domestique, d'autres préparatifs du mariage). Cette double casquette (femme, *a fortiori* jeune, non-accompagnée, loin de chez elle, travaillant et effectuant une recherche universitaire) a pu, à l'occasion, éveiller curiosité voire méfiance. Le caractère masculin de mon activité permettait mes déplacements nombreux et me conférait crédibilité et considération professionnelle. Notons que ce n'est pas le cas de toutes les chercheuses, qui parfois préfèrent ou ne peuvent interroger que les femmes (par exemple Sula Raxhimi 2015 : 64). De même, les chercheurs hommes peuvent avoir des difficultés à interviewer les femmes en tête à tête.

La réalisation de l'expérience linguistique m'a notamment fait gagner en crédibilité auprès de certains mais a pu en effrayer d'autres, car elle nécessite l'utilisation de matériel électronique (ordinateur portable, enregistreur Tascam, petit dictaphone, micro serre-tête, micro statique, avec ou sans son pied, disque dur externe…). Je déployais donc plus ou moins de matériel (en utilisant le Tascam ou le dictaphone de poche, en branchant ou non l'ordinateur et l'enregistreur sur le secteur, c'est-à-dire en multipliant ou non les câbles…) afin de rendre la situation plus ou moins formelle, donc la faire sembler plus ou moins « professionnelle ». En d'autres occasions, j'ai plutôt insisté sur mon statut d'étudiante venant apprendre la langue et la culture de la communauté – plutôt qu'explorer la réduplication de l'article défini. Cela permettait de les mettre en valeur alors qu'elles sont ignorées

par la société et les institutions – au contraire d'autres langues et cultures minoritaires reconnues et promues, comme le grec et le macédonien. Il était cocasse, aux yeux de certains, qu'une étrangère connaisse mieux leur langue minoritaire que leur langue dominante. Afin d'éviter que le travail de recherche en linguistique ne soit mal jugé ou interprété, j'ai souhaité jouer un rôle bénéfique pour la communauté par ailleurs : interprète ou auxiliaire dans des réunions (OCDE), assistante photographe.

Enfin, j'ai pu assister à un phénomène que l'on pourrait qualifier d'appropriation du chercheur par la communauté, et qui se produit notamment lorsqu'il vit au sein de celle-ci. Certains membres m'ont adoptée jusqu'à vouloir me faire parler leur variété plutôt que celle des autres, me faire vêtir ou maquiller selon leurs codes, contrôler mes affaires personnelles, éviter de me faire héberger par des personnes appartenant à d'autres groupes. J'ai souhaité, en travaillant avec plusieurs familles, ne pas risquer de devenir la chercheuse affiliée exclusivement à tel groupe ou telle famille. C'est d'autant plus important que le terrain n'est pas confiné dans un « là-bas » ou un « ailleurs » lointain : les liens continuent d'exister en France grâce au téléphone et à l'Internet – plusieurs collaborateurs demandant par exemple un soutien financier, ou encore un soutien logistique pour des membres de leur famille migrant en France. Il semble normal de continuer à rendre service aux collaborateurs qui m'ont accueillie chez eux avec une grande générosité, et dans des conditions parfois difficiles pour eux.

Sula Raxhimi (2015 : 59-60) décrit un phénomène apparenté. Son ancien emploi de « responsable des programmes d'aide au développement du Gouvernement suisse dans le domaine de la justice sociale en Albanie » lui a conféré la possibilité, les moyens et les connaissances pour faire du lobbying international en faveur des minorités. Pour cette raison, nombre de ses collaborateurs et informateurs roms, à l'exception des plus proches, ont eu du mal à comprendre et à accepter sa position de « simple » chercheuse en anthropologie. Certains, comprenant qu'elle ne pouvait leur apporter ni aide concrète ni moyen de financement, et ne voyant donc pas d'intérêt à sa recherche, lui ont pour cette raison tourné le dos.

3.4. Linguistique de données « naturalistes »

Pour les mêmes raisons que celles qui nous poussent à travailler avec un corpus plutôt qu'avec notre intuition, nous choisissons de travailler sur des phrases effectivement produites plutôt que sur des phrases construites et décontextualisées (Thuilier 2012 : 30). Des phrases énoncées dans la réalité mettent en jeu un ensemble des contraintes et des facteurs multiples dont on peut alors observer l'interaction. Néanmoins les données, en dépit de leur nom, ne sont pas un « donné » que le chercheur viendrait « cueillir » ou recueillir (collecter, obtenir…) tel quel sur le terrain, qui existerait donc en dehors de lui et des locuteurs. Elles sont le produit du travail du chercheur, qui réalise les enregistrements, choisit de collecter tel ou tel type de « texte » et de discours, mais aussi influence, parfois fortement, les locuteurs par sa présence. Le corpus, même de paroles spontanées, est un « objet fabriqué à partir de textes ou de discours produits dans des conditions "naturelles" ou "écologiques", à savoir, des conditions où les locuteurs parlent ou écrivent sans savoir que la séquence produite sera analysée à des fins linguistiques » (Thuilier 2012 : 45). Les sections méthodologiques 1 du chapitre IV, p. 201, et 1 du chapitre V, p. 287, donneront donc un aperçu détaillé de la réalité du terrain, des méthodes d'enquête et des conditions dans lesquelles a été construit puis analysé le corpus.

Tagliamonte et Baayen (2012) expliquent la difficulté de réaliser une étude à la fois sociologiquement représentative et linguistiquement suffisante. En effet, les données sont souvent disparates, les locuteurs appartenant de façon inégale aux différents groupes sociaux et linguistiques, les occurrences appartenant de façon inégale aux différents contextes linguistiques. On y trouve aussi bien de la variation inter- qu'intra-individuelle. Certains locuteurs sont au contraire catégoriques, c'est-à-dire qu'ils ne présentent aucune variation. Les contraintes agissant sur la variation linguistique sont dont simultanées, multidimensionnelles et multicausales.

De quelle manière procéder pour comprendre la structure qui fait l'objet de cette étude ? Les chercheurs travaillant sur une langue minoritaire et à tradition orale peuvent recueillir plusieurs types de données, naturelles ou construites :

a) naturelles/spontanées, issues de conversations prononcées en présence ou en l'absence du linguiste. Le corpus est alors considéré comme « authentique », « écologique » ou encore « naturaliste » – dans les limites du « paradoxe de l'observateur » (Labov 1972) ou de la conscience qu'il existe un microphone dans la pièce. La présence du chercheur et/ou du moyen d'enregistrement influence le discours et le style d'une manière profonde et difficile à mesurer : prononciation contrôlée ou plus « standard », registre plus élevé, grammaire simplifiée ou au contraire complexifiée… Pour obtenir des données non « déformées », il faudrait « supprimer » le chercheur et/ou le micro, ce qui supprimerait par là-même la collecte des données. Une solution consiste à supprimer seulement le chercheur et de cacher le microphone/la caméra (ce qui ne le supprime pas pour autant de leur conscience). À l'exception des conversations téléphoniques, on obtient alors le plus souvent des conversations de groupes, qui peuvent être difficile à entendre, à suivre voire à comprendre. D'autres facteurs peuvent influencer le discours, tels que le lieu d'enregistrement ou les témoins de la conversation, et surtout le rapport qu'entretiennent les locuteurs avec le chercheur et le statut de ce dernier (membre ou non de la communauté, intégré ou non à celle-ci, récemment ou non, etc.).

b) naturelles/spontanées mais orientées (ou dirigées), issues d'entretiens menés par le chercheur. Cela permet de contrôler le contenu de la conversation, d'obtenir certains types de discours et d'informations (texte procédurier, récit de vie, métadonnées...). Le discours est alors assez contrôlé, dans la mesure où le locuteur se sait observé et où l'enregistrement ne prétend nullement recréer des conditions réalistes de conversation. Dans le cas qui est le mien, qui ne suis pas locutrice native de la langue, il peut avoir tendance à simplifier son discours, de peur de ne pas être compris. Une solution consiste alors à demander à un membre de la communauté de mener de tels entretiens.

c) construites/spontanées, issues d'expériences linguistiques préparées et contrôlées par le chercheur. Ces données sont induites par les conditions de productions de l'expérience choisie. Cela permet de prévoir non seulement le contenu des paroles, mais aussi le vocabulaire voire les structures grammaticales employées. Le chercheur perd donc là aussi l'« authenticité » de l'énonciation car le locuteur est sur ses gardes et surveille consciemment son élocution. Le locuteur peut être seul à parler, ou bien dialogue avec le chercheur ou un autre locuteur. Si l'expérience consiste à réaliser une tâche éloignée du quotidien ou constitue une situation de communication artificielle, le locuteur peut être anxieux et sa parole s'en trouvera d'autant plus inhibée.

3.5. Linguistique de données expérimentales

Skopeteas, Féry et Asatiani (2015 : 4-5) résument comme suit les avantages et inconvénients des données issues d'une expérience linguistique :

> « *The advantage of controlled data-gathering methods is that they are maximally informative for the research question at issue: a targeted data-collection design contains the exact factors of interest crossed with the maximally relevant random factors.* […] *The generalizability of the linguistic behaviour in laboratory settings cannot be taken for granted: controlled data only contain a subset of the stylistic variation that appears in spontaneous settings, i.e., they are under-informative for the phenomena that depend on stylistic variation.* […] *Controlled data are 'idealizations' of real-world situations: speakers in the lab behave in a way in which the emotional involvement of real communication is eliminated; they avoid disfluencies that frequently occur in spontaneous speech; the target utterances are performed in isolation in order to eliminate irrelevant effects of a larger discourse* […]. *An account of spontaneous data has to overcome these obstacles by undertaking the methodological risk of 'abstraction', i.e., by averaging data varying in all these respects.* » (Skopeteas, Féry et Asatiani 2015 : 4-5)

Les données permettraient donc de généraliser les résultats et d'enquêter sur les facteurs de variation dans la structure étudiée : facteurs linguistiques bien sûr, mais également sociolinguistiques comme l'âge ou le genre (Skopeteas 2012 : 227-228). Nous pensons, nous aussi, qu'il est possible d'opérer une généralisation sur la langue à partir de données expérimentales de corpus.

> « L'analyse de données vise à déterminer des caractéristiques générales de la population langue à partir de l'échantillon dont on dispose. L'idée de chercher la généralisation à partir d'un échantillon est très répandue dans les autres sciences humaines mais a été longtemps bannie de la linguistique. […] Certaines branches de la linguistique font déjà usage des méthodes statistiques : la psycholinguistique et la sociolinguistique, sous l'influence respective des méthodes en psychologie et en sociologie. » (Thuilier 2015 : 50)

Se pose alors la question de la représentativité du corpus étudié. Le travail sur corpus représente un défi très différent en fonction de la langue d'étude. Une « grande » langue telle que l'anglais ou le français est une langue de grande extension géographique et démographique, une langue puissante et de prestige. Elle est parlée par des dizaines de millions de locuteurs, étudiée depuis des décennies par des centaines de chercheurs partout dans le monde. Elle est étayée par des siècles d'attestations écrites permettant des comparaisons diachroniques et diatopiques. Elle est d'accès facile, son apprentissage et sa culture sont soutenus par plusieurs États. Ses locuteurs souffrent rarement de discriminations à cause d'elle. On peut donc dresser un corpus représentatif et de grande taille, avec un grand nombre de locuteurs et d'occurrences. Une « petite » langue telle que le romani est une langue de grande extension géographique et démographique mais qui ne bénéficie d'aucune puissance ni prestige. Elle est parlée par un nombre moins important de locuteurs et qui vivent en petites communautés, dispersés sur l'ensemble du territoire d'un nombre important de pays, mais vivant toujours en situation de minorité linguistique. Elle est étudiée d'une manière scientifique depuis le XXème siècle seulement et de manière disparate puisqu'il existe très peu de centres de recherche qui s'y consacrent. Il existe peu d'attestations écrites anciennes permettant une profondeur historique lorsque l'on travaille en diachronie. Elle est d'accès mal aisé, son apprentissage et sa culture sont très peu soutenus

par les États et les écoles. Ses locuteurs sont le plus souvent en situation de discrimination et de grande détresse sociale et économique. Lorsque l'on travaille sur une telle langue minoritaire, on ne peut donc se fonder sur des corpus déjà établis et des bases de données existantes. Il faut construire son propre corpus lors de missions qui se déroulent toujours dans un temps restreint et sont soumises aux aléas du travail de terrain. De plus, lorsqu'aucune donnée sociologique ni statistique fiable n'est disponible sur la communauté, on ne peut qu'émettre ses propres hypothèses démographiques afin de regrouper un échantillon de la population, dont la représentativité est par conséquent subjective. Le corpus est donc d'une taille restreinte, de même que le nombre de locuteurs et d'occurrences. Même pour une langue très étudiée telle que le français, Thuilier (2015 : 94-95) écrit d'ailleurs que :

> « Nous n'avons aucune certitude sur la représentativité du corpus utilisé, car nous ne pouvons qu'émettre des hypothèses sur la diversité et la taille des documents nécessaires à la constitution d'un corpus représentatif. Étant donné que la possibilité de généraliser du corpus vers la langue repose en partie sur la représentativité du corpus étudié, nous ne pouvons qu'émettre des hypothèses sur le niveau de généralisation opérée à partir du corpus. De plus, le corpus est un ensemble d'unités linguistiques sur lequel on applique des méthodes statistiques qui nous permettent de tirer des généralités sur la population de ces unités. Le passage des généralisations sur la population des unités étudiées vers le système de la langue ne va pas de soi. Se pose alors la question de savoir si ce qui a été observé en corpus fait réellement partie de la connaissance des locuteurs. ». (Thuilier 2015 : 94-95)

3.6. Triangulation

Les différentes méthodes citées ci-dessus sont en réalité complémentaires. Skopeteas (2012 : 238-239) démontre que l'étude de corpus et l'expérimentation psycholinguistique constituent deux approches complémentaires d'un phénomène linguistique donné : expliquer le choix d'un locuteur donné d'employer telle expression linguistique dotée de propriétés particulières, dans un contexte donné. Le corpus permet d'observer le contexte dans lequel est utilisée l'expression, dans un contexte de communication réel. L'expérience permet quant à elle de contrôler les sources de la variation (interne et externe à la langue), de comparer les données entre elles et de déterminer quels facteurs discursifs influences la production. Pour l'étude d'un phénomène linguistique particulier, l'enregistrement de conversations spontanées peut se révéler insuffisant, par exemple pour un problème de registre (le phénomène n'existe qu'en registre formel) ou de fréquence (il est rare). Pour vérifier si certaines formes peuvent ou non exister, que ce soit dans les pratiques déclarées ou dans les pratiques réelles, même non-spontanées, on peut faire appel à une expérience contrôlée. C'est pourquoi il peut être pertinent de combiner les types de données (discours spontané et semi-spontané) et les approches : analyse qualitative, permise par les méthodes a) et b), et quantitative, permise par la méthode c) avec ses observations contrôlées et ses résultats comparables (Skopeteas, Féry et Asatiani 2015 : 3).

Le fait de combiner des méthodes différentes voire opposées pour étudier un même phénomène se nomme *triangulation* – terme issu d'une métaphore militaire et proposé par Denzin (1978 : 302) (cité par Jick 1979 : 602). Cela peut consister à rester dans le même type de méthodologie : par exemple dans une enquête par sondage (quantitative), à analyser le même corpus de réponses de plusieurs

manières différentes, ou encore, dans une observation participante (qualitative), à comparer plusieurs groupes de locuteurs. Cela peut consister également à croiser des méthodologies différentes, donc à combiner les méthodes quantitative et qualitative. Il s'agit d'élargir la méthodologie afin d'effectuer des recoupements d'information, de vérifier la véracité des données, de confirmer ses intuitions et hypothèses. Si une configuration ou un résultat en ressort de façon régulière, ce résultat sera d'autant plus fiable, à considérer comme valide et généralisable. Le postulat d'une telle démarche est que chaque méthode présente des faiblesses et des biais, qui seront compensés par une méthode de nature différente. Il s'agit donc d'exploiter au mieux le potentiel de chacune d'elle (Jick 1979 : 604, 608) :

- méthode qualitative : observation directe, validation personnelle de l'intuition, validation et/ou clarification par les locuteurs, « authenticité », interprétation holistique, minimisation du bruit[1] produit par les statistiques,
- méthode quantitative : systématicité, contrôle des variables et des corrélations, organisation de données complexes, reproductibilité, groupe de contrôle, minimisation du biais et de la subjectivité du chercheur, minimisation du silence[2] produit par le discours spontané, prise en compte et contrôle de la variation, donc possibilité de généralisation.

Les approches ne sont donc pas contradictoires : elles se complètent mutuellement. Selon Jick (1979 : 609), la triangulation requiert une grande part de créativité du chercheur, dans la collecte des données, leur interprétation et le croisement des résultats. Cette approche est appliquée par un nombre croissant de chercheurs travaillant sur des langues minoritaires, et même sur le romani (par exemple Adamou et Arvaniti 2011, Pădure, De Pascale et Adamou 2018, Adamou et al. 2019, Pădure 2019).

La présente étude comporte les trois types de données a), b) et c), regroupées en deux corpus :

- Les données spontanées forment le **corpus** « **spontané** ». Elles sont issues d'interviews, de narrations, de textes procéduriers (cf. section 1 p. 201). Ce corpus sert de base à une analyse qualitative du contexte de production de la structure étudiée. Une partie de ce corpus est en outre issue de mes observations sur le terrain, notamment des énoncés spontanés produits lors des interactions quotidiennes des membres de la famille qui m'a accueillie.
- Les données semi-spontanées forment le **corpus** « **semi-spontané** ». Elles sont issues d'une expérience linguistique, plus exactement de deux tâches expérimentales (cf. section 1 p. 287). Ce corpus sert de base à une analyse quantitative des constituants nominaux apparaissant dans le contexte sélectionné pour les tâches.

[1] « Le bruit, inévitable, consiste à avoir dans les résultats autre chose que la structure recherchée » (Fagard 2015 : 19).

[2] Le « silence » est l'absence, dans les résultats, de la structure recherchée. Celui-ci est courant pour les structures rares en discours spontané, notamment dans la conversation de la vie quotidienne, aux sujets aléatoires et aux situations d'énonciation très hétérogènes.

Chapitre II : Théorie

Cette partie constitue une introduction théorique aux facteurs linguistiques qui peuvent déterminer l'objet de notre étude. On examinera tout d'abord la définitude et l'article défini, avant de présenter cette partie du discours nécessaire pour que soit répété l'article défini : l'adjectif. Enfin, nous présenterons le champ d'étude de la structure de l'information, et notamment la notion de focus.

1. La définitude

1.1. Définition

Il est difficile de donner une définition universellement valable de la définitude, tant elle varie d'une langue à l'autre. Elle recouvre des fonctions linguistiques différentes, telle que l'assertion d'existence, la familiarité, l'identifiabilité, l'unicité, l'anaphore, la spécificité, la référentialité, la généricité, l'inclusiveté ou l'exclusivité (Lyons 1999 : 277). Le fait, pour un référent, d'être « défini », signifie qu'il peut être identifié par le locuteur et l'interlocuteur, par exemple parce qu'il est présent dans la situation d'énonciation, qu'il a déjà été mentionné auparavant, ou encore parce qu'il fait partie des connaissances partagées, de la culture générale commune. On peut toutefois résumer la définitude de la façon suivante : c'est lorsque le locuteur présume que l'interlocuteur a la possibilité d'identifier le référent du syntagme nominal, pour quelque raison ou par quelque moyen que ce soit.

1.1.1. Présupposition d'unicité et d'existence

La plupart des auteurs s'accordent à dire que la définitude requiert les **contraintes d'unicité et d'existence** selon lesquelles « l'information fournie par le syntagme nominal, conjointement à l'information contextuelle disponible, détermine un référent unique » (Bosch & Geurts 1990, cités par Gaiffe, Reboul et Romary 1997 : 72).

Pour que la saisie du référent fonctionne, il est nécessaire qu'il soit « unique de son espèce dans un certain contexte » (Charolles 2002 : 83). On parle de **présupposition existentielle d'unicité**, qui résiste à la fois à l'interrogation et à la négation. Cette notion implique souvent, mais pas toujours, que « l'entité dont l'existence est présupposée est déjà connue et présente à l'attention des interlocuteurs » (Charolles 2002 : 84). C'est pourquoi on dit souvent que le défini marque la **familiarité** (ou la notoriété, la notion de « connu »). L'auteur s'oppose à cette conception et montre que ce que le locuteur tient pour connu n'est pas le référent lui-même, mais « seulement le contexte particulier dans lequel les allocutaires sont appelés à comprendre ce qui rend l'entité désignée à l'aide

d'un SN défini unique de son espèce » (Charolles 2002 : 84). L'article défini ne suppose donc pas de familiarité ni de « contact préalable » des interlocuteurs avec le référent. Il ne suppose pas non plus qu'il soit au centre de l'attention, ni donné du point de vue du statut de l'information. « Au contraire, il vise à attirer leur attention sur une entité qui n'est pas encore focale dans leur esprit, mais qu'ils devraient réussir à identifier sans problème en appliquant le [sic] procédure d'identification qui lui est attachée » (Charolles 2002 : 84).

Il est important de définir ce que l'on entend par « contexte » ou « information contextuelle disponible ». Selon Gaiffe, Reboul et Romary (1997 : 86), cette notion doit s'entendre au sens le plus large, « comme une situation, un domaine, un espace, etc. à l'intérieur duquel on peut identifier le référent recherché et où il y a des objets d'autres types ». Pour que ce principe soit universellement valable, l'information contextuelle est une information co-textuelle (le discours qui précède et qui suit immédiatement) mais aussi context-uelle (connaissance du monde, situation d'énonciation).

- Le co-texte peut fournir soit le référent lui-même, repris ultérieurement par l'expression définie, soit un autre référent permettant une anaphore associative (« *Nous entrâmes dans un village. L'église était située sur une hauteur.* ») (Gaiffe, Reboul et Romary 1997 : 85).
- Le contexte, quant à lui, fournit les informations de la culture ou de la situation d'énonciation qui sont nécessaires à l'identification du référent. « Contrairement à ce que peut laisser croire l'analyse sur des textes, il est extrêmement fréquent qu'un groupe nominal défini construise un référent du discours. Ainsi, à table on peut demander "le sel" même si personne n'a précédemment parlé "d'un sel" » (Gaiffe, Reboul et Romary 1997 : 85). Autre exemple, emprunté à Charolles (2002 : 92) : Lors d'un repas, une personne tend un plateau de fromage à un convive, qui refuse et répond « J'attends le gâteau » – sans avoir vu ou avoir entendu parler d'un gâteau. Sa connaissance du monde l'informe qu'étant donné ce genre de repas, un dessert doit forcément venir après le fromage, et le dessert est le plus souvent un gâteau. En effet, « les SN définis sont très souvent utilisés pour désigner une entité qui se trouve dans l'environnement perceptif des participants et dont l'existence et l'unicité sont notoires » (Charolles 2002 : 91). Même si le référent n'est pas donné en discours ni dans la situation, et n'est pas directement accessible, l'anaphore associative le rend inférable du fait que la situation entretient une relation régulière avec lui. « La situation rend mentalement accessible une entité unique qui lui est stéréotypiquement associée » (Charolles 2002 : 92).

Gaiffe, Reboul et Romary (1997) en déduisent que « la description[1] définie p[eut] créer un référent de discours », ce « par rapport à un contexte pré-existant » (Gaiffe, Reboul et Romary 1997 : 85). On ne devrait donc pas (ou pas seulement) parler d'une présupposition d'existence du référent, car l'interprétation d'une expression définie « ne présuppose pas l'existence d'une représentation du référent avant l'apparition de la description définie concernée » (Gaiffe, Reboul et Romary 1997 : 89). Mieux vaut parler d'une **présupposition de pré-existence du contexte** où pourra être créé le référent du discours défini (Gaiffe, Reboul et Romary 1997 : 85, 86, 89).

[1] l'expression, le syntagme

1.1.2. Expressions définies complètes et incomplètes

Charolles (2002 : chap. 6) distingue les descriptions ou **expressions définies complètes** de celles **incomplètes**. « Les SN définis signalent qu'une entité ou un ensemble d'entités est la seule ou sont les seules d'un certain type. Cette indication (présupposition) a besoin d'être justifiée contextuellement. Soit le SN défini comporte en lui-même des adjoints qui spécifient les circonstances dans lesquelles cette unicité est garantie (descriptions définies complètes), soit le SN défini ne fournit aucune information de nature à satisfaire cette exigence (descriptions définies incomplètes) » (Charolles 2002 : 103).

- Un syntagme comme *la championne de France 1999 de planche à voile* ou *la voiture immatriculée 748 KNB 75* n'est valide que pour un seul référent. On peut dire que la définitude y est complète en ce sens qu'elle permet d'identifier sans erreur le référent. Ce type d'expression est constitué d'un déterminant défini, du nom et d'un ou de plusieurs modifieurs « qui restreignent sa dénotation au point que l'expression obtenue ne vaut, comme un Np, que pour une seule entité ou un seul groupe d'entités » (Charolles 2002 : 76): adjectif, superlatif, complément prépositionnel, proposition relative déterminative, adjoint de nature nominale (numéro, date… incluant des indications spatio-temporelles ou personnelles spécifiant le contexte de validité).

Une telle expression se caractérise par son autonomie référentielle : elle permet d'identifier le référent indépendamment de toute information sur le contexte où elle est employée. Ceci inclut les noms propres, les noms communs uniques (baptisés *unica* et pour lesquels l'orthographe hésite souvent à mettre une majuscule : *la lune, le soleil, l'équateur, le nord*), les noms massifs, génériques et abstraits « dont la référence est établie de manière autonome et qui fonctionnent sans problème comme des SN définis complets », les phénomènes climatiques et naturels (*le vent, la pluie, la foudre, la lumière*), les types désignant « des entités intrinsèquement uniques parce qu'existant indépendamment de leurs incarnations particulières » (*la lettre A, le chiffre 4, l'as de pique*), et enfin les expressions métalinguistiques (*le mot "soif"*) (Charolles 2002 : 76-77). En revanche, cela exclut les expressions déictiques (dépendantes de la situation d'énonciation) et les expressions renvoyant à un référent unique mais dépendant du moment et du lieu de l'énonciation : noms de rôle, titre et fonction sociale, grades, superlatifs relatifs (Charolles 2002 : 76). On peut parler d'incarnation d'un rôle unique et stéréotypé, pour l'identification duquel l'article défini suffit : « la description ne peut être satisfaite que par une personne unique, l'indication du titre qu'elle est la seule à porter est un moyen commode pour en parler » (Charolles 2002 : 94).

- Un syntagme comme *l'entrée du magasin* ou *le premier ministre* ne sont pas valides pour un seul référent. On peut dire que la définitude y est incomplète en ce sens qu'elle ne peut assurer totalement l'identification. Ce type d'expression est constitué des mêmes éléments qu'une expression définie complète. Elle peut être syntaxiquement aussi complexe et sémantiquement aussi spécifique qu'une expression définie complète (Charolles 2002 : 77).

Une telle expression se caractérise par son absence d'autonomie référentielle : elle ne permet pas

d'identifier le référent sans contexte ni information présupposée. Le mode d'établissement de la référence varie d'une expression à l'autre (Charolles 2002 : 77-78). Ces expressions sont fort courantes dans la vie quotidienne et n'empêchent nullement l'intercompréhension : notre connaissance du monde nous permet de savoir à quoi renvoient *la clé, le chat, les plantes, l'infirmière* lorsqu'il s'agit d'éléments de notre entourage. Notre connaissance de la situation d'énonciation nous permet de savoir à quoi renvoie *la dame* lorsque l'on est en train de jouer aux cartes et que la dame de trèfle est sur la table. Ces éléments font partie du « contexte restreint où l'expression référentielle utilisée est interprétée » (Charolles 2002 : 78).

Charolles (2002 : 87) écrit que l'article défini suggère une expérience partagée ou une histoire commune ». Il est nécessaire que l'ensemble des participants de la communication aient accès à ce même contexte restreint, à cette expérience ou cette histoire, pour que le syntagme défini incomplet n'échoue pas à permettre l'identification et l'intercompréhension. L'auteur rejoint ici la théorie de la présupposition de pré-existence du contexte de Gaiffe, Reboul et Romary (1997). L'usage de la définitude incomplète sollicite donc les « capacités de catégorisation et de discrimination catégorielle des participants à la communication » et « le contact mental […] que ceux-ci peuvent avoir avec les ingrédients de la situation qui est au centre de leur intérêt » (Charolles 2002 : 80-81). Plus largement, on peut dire que l'usage d'une telle expression révèle la façon dont le locuteur conçoit le référent et l'ensemble des référents « rivaux » qui peuvent faire partie de la dénotation de l'expression (Charolles 2002 : 80). Il révèle également la capacité d'interprétation qu'il attribue à l'interlocuteur : le locuteur présuppose que son interlocuteur soit capable d'interpréter le syntagme, de repérer le référent comme unique et de l'identifier. « Cela revient à dire que ce que présuppose une description définie incomplète ce n'est jamais que la présomption que son énonciation est justifiée dans la situation où elle est produite » (Charolles 2002 : 85). « Dans tous les cas, le locuteur qui recourt à un SN défini présume que celles et ceux à qui ils s'adressent [sic] vont être à même de récupérer sans difficultés les raisons faisant qu'on leur parle d'une entité comme unique de son espèce, d'où des effets de connivence qui sont très sensibles à l'intuition » (Charolles 2002 : 103-104).

1.1.3. La catégorie grammaticale

La définitude est une catégorie du discours distincte du fait, pour une langue, de présenter ou non un article défini : la première ne présuppose pas le second, même si le second présuppose la première. Il s'agit d'une catégorie d'ordre sémantique et pragmatique qui peut en comprendre d'autres, selon les auteurs, le fait d'être spécifique/particulier/individuel, familier, situationnel, connaissance générale du monde, anaphorique, référence transversale ou associative, identifiable, unique, inclusif dans le contexte (Lyons 1999 : chap. 1). Par « définitude », on entend également une catégorie grammaticale, qui n'est donc pas présente ou exprimée dans toutes les langues. Lorsque la définitude est marquée explicitement dans une langue donnée, elle est selon Lyons (1999 : 277) issue de la grammaticalisation de la définitude sémantique ou pragmatique. On peut dire que la définitude grammaticale est la définitude sémantique/pragmatique grammaticalisée. En fonction des langues, le trait de la définitude exprime un ou plusieurs des catégories pragmatiques citées : selon Lyons (1999 : 278), celle qui est principlement grammaticalisée est l'identifiabilité. Cette catégorie universelle qui permet d'interpréter le discours n'est pas incarnée par un élément grammatical dans toutes les

langues ; lorsqu'elle l'est, c'est prototypiquement par la définitude. Cela n'empêche pas certaines instances de l'identificabilité, par exemple la généricité, d'être exprimées par autre chose que par la définitude dans certaines langues. Cela n'empêche non plus pas la définitude d'exprimer d'autres catégories, telles que l'inclusivité selon Lyons (1999 : 278). La variation est très importante, mais le socle des usages et communs est très important, ce qui permet de concevoir et d'étudier cette catégorie.

Elle est souvent conçue en opposition à l'indéfinitude, notamment dans les langues opposant un article défini à un article indéfini. En grammaire, le défini signifie que l'interlocuteur peut identifier le référent du syntagme, en raison de la spécificité du nom, grâce à ses modifieurs, par le contexte ou encore par sa connaissance du monde. « Le défini signifie simplement que le lecteur en sait assez sur ce dont il est question : il est en quelque sorte plongé immédiatement dans l'univers de référence du texte » (Poitou 2010). Au contraire, « l'indéfini indique que le référent du syntagme n'est pas identifié à l'intérieur de la catégorie représentée par les lexèmes du syntagme » (Poitou 2010). Cependant, le défini comme l'indéfini, au singulier comme au pluriel, peuvent avoir une valeur de générique. On entend par générique la référence à une classe entière, ou encore la généralisation que l'on tire d'une classe entière (Lyons 1999 : 179). C'est pourquoi, plutôt que d'opposer binairement un trait [+défini] à un trait [-défini], il nous semble plus pertinent de concevoir la définitude comme un continuum allant de l'absence de détermination au démonstratif (Figure 6).

[ø ---------- IND ---------- DEF ---------- DEM]

Figure 6 : Continuum de la définitude

Le démonstratif, dans son usage déictique, fait référence à la situation d'énonciation, en localisant le référent dans le temps ou l'espace en fonction de sa proximité à la 1ère, 2ème ou 3ème personne. Cet usage peut être étendu au discours lui-même, usage que l'on nommera anaphorique. L'article défini provient de tels usages. Ce continuum s'exprime par différents moyens qui peuvent se combiner :

- suprasegmentaux : place et longueur de l'accent tonique, ton, intonation…
- morphologiques : déterminant, article libre ou lié, désinence flexionnelle sur le verbe ou sur le nom…
- syntaxiques : réduplication, ordre des mots…

Lyons (1999 : 49-50) note que certaines langues expriment morphologiquement la seule définitude (irlandais), d'autres la seule indéfinitude (turc), et d'autres la définitude aussi bien que l'indéfinitude (danois). Selon Lyons (1999 : 48-49), le fait d'encoder la définitude, notamment par un moyen morphologique et par un défini simple (section 1.3.1, p. 64), est un trait aréal plus que génétique. Une zone particulièrement riche en langues exprimant la définitude est le bassin méditerranéen et l'Europe occidentale, avec des langues qui ont essaimé, via les colonisations, en Afrique et en Amérique centrale et méridionale. Le lieu d'expression privilégié de la définitude dans une phrase est le syntagme nominal (sur le nom, un déterminant, un adjectif) mais pas exclusivement : en hongrois, le verbe se fléchit différemment selon l'(in-)définitude de son objet direct (Moravcsik 2003 : 400-405,

Lyons 1999 : 86). C'est le cas dans plusieurs langues ouraliennes (Lyons 1999 : 86). Dans d'autres langues encore, comme le chinois, la définitude se marque par un changement de l'ordre des mots (Lyons 1999 : 88-89).

1.2. Effet sur la structure de l'information

Le marquage de la définitude a pour fonction d'aider l'interlocuteur à comprendre le statut des référents au sein de l'énoncé et du discours. Il entre donc dans une stratégie de structure de l'information. La définitude a souvent une fonction anaphorique et encode donc plutôt l'information donnée que nouvelle. Lyons (1999 : 227-236) explique les effets de la définitude du point de vue de la structure du discours, c'est-à-dire de la manière dont le message est présenté à l'interlocuteur. Le message est mis en lien avec le contexte d'énonciation, les autres messages précédemment délivrés, ceux que l'on s'apprête à délivrer ou auxquels l'interlocuteur peut s'attendre – pour plus d'informations, on pourra consulter (Erteschik-Shir 2014).

La définitude a souvent une fonction anaphorique et encode donc plutôt l'information donnée que nouvelle. Si elle exprime l'identifiabilité (l'anaphorique, la chose incluse dans le co-texte) ou la familiarité (le situationnel, la connaissance du monde, la chose incluse dans le contexte), elle a de grandes chances d'être présente dans dans les syntagmes exprimant de l'information donnée (Lyons 1999 : 233). Notons toutefois que définitude et information donnée ne coïncident pas complètement, car un syntagme défini peut apporter une information nouvelle (exemples (12) ou (13), pour lesquels on peut discuter si les référents font partie de la connaissance du monde et sont donc en réalité accessibles)… et un indéfini peut apporter une information donnée (Lyons 1999 : 232).

La définitude contribue dans ce cas à marquer le topique – ce qui n'empêche pas le topique de pouvoir être indéfini (Erteschik-Shir 2014 : 26). Ainsi sert-elle à guider l'interlocuteur à comprendre la structure informationnelle de la phrase. Voilà pourquoi, selon Lyons (1999 : 233), les langues qui ont des marqueurs de topique (coréen, japonais) n'ont souvent pas de marqueur de définitude (français, anglais, arabe), à quelques exceptions près (tzotzil).

1.3. L'article défini

1.3.1. Définition

L'article défini sert, selon les langues, à encoder la définitude et/ou la spécificité (De Mulder et Carlier 2011 : 524). Il est plus répandu, dans les langues du monde, que l'article indéfini : une langue possédant un article indéfini grammaticalisé a plus de chances d'avoir également un article défini que l'inverse (De Mulder et Carlier 2011 : 524) – le turc, langue avec article indéfini sans article défini, constitue ici une exception typologique. La classe particulière d'éléments nommée *article* est un type de déterminant – du moins suivant la tradition grammaticale indo-européenne. L'article défini est le

porteur par excellence du trait [défini] – contrairement au déterminant ou au possessif, qui l'associent à d'autres traits. On parlera donc, suivant la terminologie de Lyons (1999), de défini simple (*simple definite*) et de définis complexes (*complexe definites*). Dryer (2013) propose la définition suivante : morphème qui accompagne le nom et qui code la définitude ou la spécificité (section 1.3.2, p. 65). Cela permet d'inclure dans la définition des affixes et des démonstratifs utilisés pour marquer la définitude.

Kyriakaki (2014a : 260) et Kyriakaki (2014b) propose quant à elle une typologie tripartite des articles définis. La définitude consiste pour elle en deux traits : la familiarité et l'unicité. Elle fait le lien entre trois degrés de définitude, exprimant la familiarité l'unicité, la familiarité seule, ou aucun de ces deux traits. Il existe donc trois types de déterminants :

- pleins : entièrement spécifiés pour la définitude, ils expriment à la fois l'unicité et la familiarité (anglais, allemand du nord).
- partiels : sous-spécifiés pour la définitude, ils expriment le seul trait de la familiarité. Ils introduisent des noms propres, des noms génériques, voire des modifieurs restrictifs (grec, allemand, français).
- zéro : explétifs, ils n'expriment aucun des deux traits (italien).

Dans les langues indo-européennes, à l'exception du grec, il s'agit d'un phénomène récent. L'article n'est pas hérité de l'indo-européen, et pourrait même, dans les langues parlées dans le bassin méditerranéen, avoir pour origine une influence des langues sémitiques (De Mulder et Carlier 2011 : 522). L'article défini provient bien souvent d'un démonstratif, distal ou de troisième personne, ayant progressivement perdu tout ou partie de son pouvoir déictique jusqu'à ne plus exprimer que la définitude (Figure 7). Il a pu par la suite devenir indispensable à l'actualisation du nom dans le syntagme nominal : l'article défini a pour caractéristique d'accompagner le nom de manière quasi systématique dans les langues où il existe. On peut parler dans ce cas de grammaticalisation, un processus par lequel un morphème lexical devient un morphème grammatical, en perdant son sens plein par un processus de *semantic bleaching*, « délavement sémantique », jusqu'à n'exprimer plus qu'un concept grammatical (Lyons 1999 : 275). Le démonstratif est certes déjà un morphème grammatical, mais il est sémantiquement plus chargé que l'article défini. Il s'agit pour Lyons (1999 : 331) d'un « processus diachronique très général de dévaluation du contenu lexical » : le démonstratif perd des traits sémantiques, ce qui le fait glisser sur le continuum de la Figure 6. La grammaticalisation s'accompagne le plus souvent d'un processus de réduction (morpho-phonologique : l'article défini est souvent monosyllabique et ne peut, dans la plupart des langues, porter l'accent tonique (Lyons 1999 : 64).

1.3.2. Typologie

L'article défini apparaît dans une position fixe par rapport au nom-tête, et peut avoir différent statuts : déterminant ou affixe. Il marque alors le seul nom ou bien l'ensemble du syntagme nominal. Dryer (2013) propose une typologie des articles définis dans les langues du monde :

1) mot défini distinct du démonstratif, éventuellement identique à un pronom de troisième personne (216 langues recensées),
2) mot démonstratif utilisé comme marqueur de définitude, à la syntaxe éventuellement différente selon qu'il est utilisé comme démonstratif ou article défini (69 langues),
3) affixe défini sur le nom et/ou un autre mot du syntagme nominal (92 langues),
4) pas d'article défini, mais présence d'un article indéfini dans la langue (45 langues) – zéro peut alors exprimer soit la définitude, soit la définitude ou l'indéfinitude en fonction du contexte,
5) ni article défini ni article indéfini (198 langues).

Certaines langues expriment la définitude d'une seule manière (le français présente un article libre antéposé au nom, de type 1), d'autres de plusieurs manières (le danois présente un article libre antéposé au nom, de type 1, et un article lié postposé au nom, de type 3, qui n'apparaît qu'en présence d'un modifieur). Notons qu'il n'existe pas de lien entre défini et indéfini : une langue peut présenter un type d'article défini et d'article indéfini distincts, et ces éléments peuvent avoir une syntaxe elle aussi distincte (Lyons 1999 : 106). L'anglais présente un article défini et un article indéfini qui sont tous deux des morphèmes libres antéposés au nom ; le lakhota présente deux morphèmes libres postposés… au contraire l'amharique et le roumain ont tous deux un article indéfini libre et antéposé au nom, mais un article défini lié. Lyons (1999 : 49) propose de distinguer entre trois types de langues : celles qui comportent un article défini mais pas d'article indéfini (grec ancien, irlandais), celles qui comportent un article indéfini mais pas d'article défini (turc, domari[1]), celles enfin qui comportent les deux à la fois (français, danois). Selon De Mulder et Carlier (2011 : 523-524) au contraire, l'article défini est en fait plus répandu que l'article indéfini dans les langues du monde, à tel point que l'article indéfini naît souvent dans des langues où existe déjà un article défini, et il naît également plus tôt.

Dryer (2013) constate que les langues à article défini distinct des démonstratifs (type 1) sont présentes surtout en Europe occidentale, dans tout le centre de l'Afrique de l'est à l'ouest, dans le Pacifique et en Méso-Amérique. Le romani en fait partie. Il s'agit donc d'un phénomène aréal. Il existe un continuum selon Schroeder (2006) cité par De Mulder et Carlier (2011 : 522) : à l'ouest de l'Europe, l'article défini est distinct du démonstratif (anglais, français), au centre de l'Europe, il a la même forme que celui-ci (allemand), à l'est de l'Europe il n'y a pas d'article défini mais un démonstratif (russe), qui peut éventuellement être utilisé, dans des contextes restreints, comme un article (finnois).

Lorsqu'il est un mot distinct du démonstratif, l'article défini occupe une position adnominale fixe. Il peut être un morphème libre, comme en français (*le*), anglais (*the*) ou irlandais (*an*), ainsi qu'en romani (*o*) ou bien un morphème lié, comme en albanais (-*i*), arménien (-*ə/n*) ou abkhaze (*a*-) (Lyons 1999 : 70). L'auteur classe aussi les langues en fonction de la position de l'article : antéposé au nom (français, anglais, irlandais, abkhaze), postposé à celui-ci (abkhaze, albanais), ou bien encore cliticisé à l'ensemble du syntagme nominal en une position fixe (par exemple en position de Wackernagel, deuxième place du syntagme nominal, quels que soient le premier élément et la place du nom :

[1] En domari, le marquage différentiel de l'objet encode toutefois la définitude (Herin 2014) – alors qu'en romani il encode l'animéité.

amharique *-u*, bulgare *-ət*, roumain *-ul*, basque *-a*). On peut donc dire que l'article défini affixal est soit un marqueur de nom, soit un marqueur de syntagme nominal (De Mulder et Carlier 2011 : 524).

1.3.3. Fonctions en diachronie

Greenberg (1978) montre que l'article défini est issu de la grammaticalisation du démonstratif en suivant le processus de la Figure 7 :

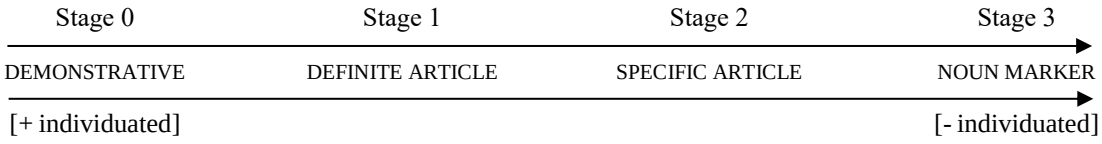


Figure 7 : Processus de grammaticalisation de l'article défini (De Mulder et Carlier 2011 : 524)

Le démonstratif proximal indique que le référent peut s'identifier pleinement dans le contexte immédiat de l'énonciation. Ce n'est pas le cas du démonstratif distal, qui peut activer une connaissance particulière partagée par le locuteur et l'interlocuteur pour effectuer l'identification. On observe le passage d'un morphème effectuant directement la référence (démonstratif) à un morphème l'effectuant indirectement (article défini, par exemple dans l'anaphore associative, et le recours à la connaissance générale du monde). De Mulder et Carlier (2011 : 526) font l'hypothèse que la naissance de l'article défini provient d'un besoin de focaliser l'attention sur un référent important, qui sinon n'en ferait pas l'objet. Dans une langue à détermination zéro, l'usage d'un démonstratif, ou d'un déterminant en général, a un impact très important. Dans un premier temps, cet usage implique un renforcement pragmatique, et c'est la généralisation de cet usage qui entraîne ensuite un affaiblissement de la force déictique ou démonstrative du déterminant. L'article défini en lui-même a perdu l'ancrage dans la situation d'énonciation : les interlocuteurs partagent une connaissance du monde et du discours.

De Mulder et Carlier (2011 : 527-528), s'appuyant sur Löbner (1985), montrent qu'historiquement, l'article défini possède d'abord un sens de **définitude pragmatique**. Cet usage, qu'il partage avec le démonstratif distal, permet d'identifier le référent dans le contexte particulier de l'énoncé : par exemple usage déictique (le référent est présent dans la situation d'énonciation), usage anaphorique (le référent a été mentionné en discours), usage cataphorique (introduisant une proposition relative). Il prend ensuite un sens de **définitude sémantique**, plus grammaticalisée. Le référent de l'expression peut s'identifier indépendamment du contexte : par exemple usage situatif (le référent peut se déduire de la situation d'énonciation), anaphore associative, référence à une entité unique dans l'univers, référence générique. De même que le processus menant du démonstratif à l'article défini est un continuum et non une série d'étapes discrètes, le processus menant de l'usage défini pragmatique à l'usage défini sémantique est un continuum :

> « *The distal demonstrative becomes a definite article when the anchorage in the speech situation is lost and the use of article no longer requires specific knowledge shared by speaker and hearer to be activated to identify the referent of the noun phrase. Rather, the definite article conveys the*

instruction that the descriptive content of the noun allows the identification of the referent in a univocal way, in virtue of its structural links with a frame of accessible knowledge, these links being often of a stereotypical nature. » (De Mulder et Carlier 2011 : 527-528)

Si le processus de grammaticalisation se poursuit, l'article défini acquiert un sens spécifique, défini puis indéfini (ainsi l'article *ang* en tagalog). Dans ce type de contexte, certaines langues du monde utiliseraient plutôt un article indéfini, notamment lorsqu'il s'agit d'introduire un référent nouveau (De Mulder et Carlier 2011 : 528). Le nombre de contextes d'emplois possibles s'étend, en raison de l'érosion sémantique subie : l'article défini ne permet plus nécessairement d'identifier le référent. L'article spécifique ne nécessite ni ancrage, ni connaissance commune, ni même l'identification du référent : il n'y a plus définitude (ni pragmatique, ni sémantique).

Selon Greenberg (1978 : 69), le développement normal d'un article devenu spécifique est de devenir « un simple marqueur de nominalité, un marqueur de classe nominale ou de genre, voire un morphème de cas », les affixes de genre dérivant d'anciens articles définis dans de nombreuses langues (Lyons 1999 : 337). Il n'est plus qu'un marqueur grammatical, de nominalité, de genre, qui actualise le nom dans la phrase. Son rôle est morphologique et syntaxique, sans apport sémantique.

1.3.4. Fonctions en synchronie

L'article défini a selon Dryer (2013) deux fonctions : faire référence à une chose préalablement mentionnée en discours, et supposée connue de l'interlocuteur (par la situation d'énonciation, la connaissance générale du monde, l'inférence du contexte ou du co-texte…). Il nomme la première fonction anaphorique, la seconde non-anaphorique.

Lorsqu'il existe dans une langue, l'article défini sert en outre à construire le superlatif des adjectifs : *la moto la plus rapide*, *the cleverest girl*, la fille la plus intelligente. Selon Lyons (1999 : 246), la définitude est ici marquée « dans le sens étroit, le sens grammatical », et non dans un sens plus large, d'une définitude sémantique ou pragmatique. Il note qu'un syntagme nominal comprenant un adjectif au superlatif ne peut être introduit que par un article défini (*la moto la plus rapide*) ou par un possessif (*ma moto la plus rapide*), mais non par un autre déterminant, défini complexe ou indéfini (**une moto la plus rapide*, **cette moto la plus rapide*, **chaque moto la plus rapide*, **plusieurs motos les plus rapides*). On peut expliquer la présence de l'article défini par l'unicité que suppose le superlatif : le référent (la moto) a une propriété (la rapidité) à un point que les autres référents n'ont pas (aucune moto n'est aussi rapide que celle dont il est question). Pour une étude approfondie sur le superlatif, on pourra consulter Hawkins (1978) ainsi que l'analyse d'Alexiadou (2014 : 68-75, 112).

Lyons (1999 : 337) observe que l'usage de l'article défini peut être étendu au-delà de celui de la définitude. L'article défini peut présenter simultanément les fonctions suivantes :

1. Anglais : défini simple
2. Français : défini simple, générique
3. Italien : défini simple, générique, possessif
4. Grec : défini simple, générique, possessif[1], nom propre

On entend par générique un syntagme faisant référence à une classe entière ou transmettant une généralisation sur l'ensemble d'une classe – qui comprend toutes les entités satisfaisant sa définition (Lyons 1999 : 179). Pour le français, De Mulder et Carlier (2011 : 523-524) proposent une répartition des usages de l'article défini et de l'article dit partitif selon une échelle de l'individuation : plus le référent est hautement individué, plus l'usage de l'article est historiquement ancien et grammaticalement systématique ; moins le référent est individué, spécifique et concret, plus la présence de l'article est historiquement récente et grammaticalement fluctuante. « *The emergence of the different articles and the progressive spread to new contexts of use can be accounted for in terms of the scale of individuation* » (De Mulder et Carlier 2011 : 524), représentée par la Figure 8.

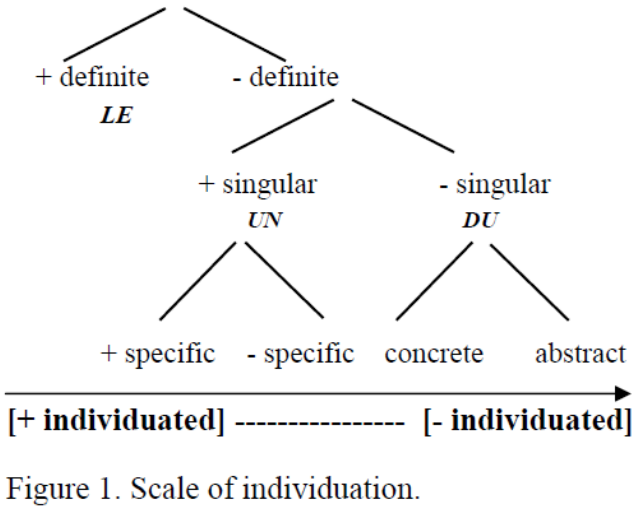


Figure 8 : Échelle d'individuation (De Mulder et Carlier 2011 : 524)

En français, la généralisation de l'article défini à tous les noms communs fait de l'article défini un simple marqueur de nom indiquant la frontière gauche du syntagme nominal.[2] Les contextes d'emploi de l'article se sont bel et bien étendus jusqu'à en faire un marqueur de syntagme nominal, et à revêtir la fonction morpho-syntaxique plutôt que sémantique (De Mulder et Carlier 2011 : 524).

[1] Voir Alexiadou (2005) pour une étude de la définitude et de la possession en grec moderne.

[2] De Mulder et Carlier (2011 : 524) relèvent un « domaine résiduel » où le nom commun se passe, en français, d'article défini : les noms propres (déterminés intrinsèquement), le vocatif (ancré dans la situation d'énonciation), l'introduction par d'autres déterminants (possessifs, démonstratifs…). Il existe en réalité d'autres contextes, listés exhaustivement par Riegel, Pellat et Rioul (2009 : 163-167), expliqués et motivés par Lebas-Fraczak (2015 : 72-82).

1.3.5. Usage avec les noms propres

Dans un premier type de langues présentant un article défini, les noms propres peuvent en être accompagnés (Lyons 1999 : 122), de manière variable en fonction de raisons d'ordre :

- grammatical : l'arménien occidental emploie l'article défini avec certains cas (accusatif, datif, ablatif) mais pas d'autres (nominatif, instrumental, le génitif à l'écrit),
- pragmatique : l'allemand emploie l'article défini devant les prénoms des femmes et des hommes en registre familier mais non soutenu ; devant les noms de famille des femmes en registre courant ; devant les noms de famille des femmes et des hommes en registre familier lorsque la phrase a une portée émotionnelle (offense ou compliment)[1] ; devant les noms de famille mis au pluriel afin de désigner plusieurs membres en registre courant.

La forme de l'article défini peut varier selon que le nom est commun ou propre : le catalan emploie l'article de manière optionnelle sur les noms propres, la forme du masculin singulier différant lorsque le nom propre commence par une consonne (Lyons 1999 : 123

Dans un premier type de langues présentant un article défini, les noms propres peuvent aussi ne pas être accompagnés d'un article défini. C'est le cas du français pour la plupart des noms propres – mais pas tous (*la Corse*, *l'ONU*) et de l'anglais, qui peut toutefois déterminer le nom propre par un démonstratif ou un possessif : *this Peter, my Annie* (Lyons 1999 : 122). Les référents des noms propres étant censés être uniques, ils ne nécessitent pas de marquage de la définitude (Lyons 1999 : 193). Le nom propre n'est pas censé dépendre du contexte : il désigne une entité dont la référence est unique et stable, quel que soit le l'énonciateur et les conditions d'énonciation. Le nom propre est toujours spécifique et individuel, contrairement au nom commun qui peut être spécifique ou générique, individuel ou collectif. C'est pourquoi un nom propre employé comme nom propre n'a pas besoin, en principe d'être articulé.

Comment expliquer l'usage de l'article défini dans un autre type de langues, dont ressortent le grec, l'albanais et le romani : celles qui ont généralisé l'usage de l'article défini des noms communs aux noms propres, notamment anthroponymes et toponymes ? Pour une présentation détaillée du nom propre articulé en grec, on pourra consulter Marinis (2003 : chap. 7). Lyons (1999 : 194-195) fournit une explication. Une première raison est que les noms propres ont parfois un référent non unique, à l'instar des noms communs : plusieurs personnes ont pu s'appeler Cicéron ou Michael Jordan dans l'histoire du monde et il existe de nombreux référents possibles à *la/une Guinness* (dénombrable), *du Destop* (indénombrable) – d'une manière générale, les noms propres d'objets de consommation[2] peuvent être définis aussi bien qu'indéfinis. Une seconde raison est que ces langues ont entrepris un processus de *recatégorisation*, en mettant les noms communs, massifs et indénombrables sur le même plan. Cela permet d'utiliser un nom d'un certain type comme un nom d'un autre type (Lyons 1999 :

[1] Certains parlers du français et de l'italien présentent un usage proche de l'allemand.

[2] On peut considérer qu'il s'agit d'une antonomase, d'où un comportement morphologique et syntaxique de nom commun : déterminant requis, pluriel possible.

196). Le romani et le grec suivent une règle cohérente : articuler obligatoirement tous noms, quel que soit leur type (commun ou propre, dénombrable ou non, générique ou spécifique). Le français, lui, requiert l'article défini pour les noms génériques (au singulier comme au pluriel) et les noms indénombrables (Lyons 1999 : 197), et partiellement pour les noms propres (pas avec *Mélusine* ni *Berlin*, mais avec *la Russie* ou *l'Assemblée Nationale*). Au contraire, l'anglais n'emploie d'article défini ni avec les noms génériques, ni avec les indénombrables, ni avec les noms propres (Lyons 1999 : 197). Winter (2004 : section 5) propose une analyse détaillée du fonctionnement sémantique des noms propres. Il se rapproche de celui des prédicats nominaux nus. En anglais, les noms propres présentent un article défini vide, et sont soumis à l'exigence d'unicité (*uniqueness requirement*). Ils sont des noms nus possédant une structure lexicale supplémentaire permettant leur syntaxe (absence d'article) et leur sémantique (dénotation d'une entité, référence « unique ») particulières. Il montre que les titres, tels que l'anglais *president*, se comportent comme des noms propres et possèdent les mêmes propriétés syntaxiques et sémantiques. C'est le cas également en grec.

Notons que le type grec/albanais/romani connaît des exceptions. Le nom propre peut ne pas être articulé : en albanais lorsque le nom suit une préposition de lieu, est en apposition ou au vocatif, en grec lorsque le nom propre revêt la fonction d'attribut du sujet, comme s'il s'agissait d'une propriété (Lekakou et Szendrői 2014 : 232-235).

2. L'adjectif

2.1. Définition

On entend par *adjectif* un morphème lexical libre qui modifie le nom. Il est difficile voire impossible de définir la classe des adjectifs (Goes 1999 : 11) dans les langues du monde, tant les critères sont nombreux et variables d'une langue à l'autre.

Le critère de la grammaire traditionnelle est **sémantique** : l'adjectif désigne une propriété, une caractéristique du référent du nom. Il caractérise le nom de manière restrictive (ou non) selon qu'il en réduit (ou non) l'extension. En y incluant un critère syntaxique, Gaatone (2016 : 16) propose la définition suivante : « sera considéré comme A tout mot susceptible d'accompagner facultativement[1] un SUB ou un pronom (PRO) à l'intérieur d'un syntagme nominal (SN), qui s'accorde avec lui en genre et en nombre, mais qui est inapte à former un SN à lui seul, avec son support, à quelques exceptions près ». L'adjectif peut alors être accompagné d'un complément ou bien être modifié par un adverbe, l'ensemble formant un syntagme adjectival.

Un autre critère peut être **morphologique** : « du point de vue intuitif une langue possède la classe lexicale des adjectifs quand les adjectifs peuvent être distingués des noms et des verbes » (Matushansky 2005 : 32). « Ainsi dans les langues romanes, sémitiques et slaves, les adjectifs peuvent être distingués par le fait qu'ils s'accordent en genre et nombre, tandis qu'en anglais, en allemand et dans d'autres langues germaniques, les adjectifs peuvent se combiner avec le morphème lié de comparaison (-*er*) et des adverbes comme *très* » (Matushansky 2005 : 32). C'est le cas du romani. Mais toutes les langues ne présentent pas de propriétés morphologiques distinctives.

Le critère peut être **syntaxique** : *« First, adjectives can be direct attributive modifiers of nouns, but nouns and verbs cannot be.* […] *Second, adjectives can be the complements of degree heads like* so, as, too, *and* how *in English, but neither nominal nor verbal projections can be. Third, adjectives can be resultative secondary predicates, unlike nouns and verbs* » (Baker 2003 : 191). L'adjectif est donc la catégorie qui ne présente ni des propriétés nominales, ni des propriétés verbales. Le premier et le troisième critère sont purement syntaxiques : la possibilité de fonctionner comme une épithète et en tant que résultatif. Le troisième critère, la compatibilité avec des morphèmes de degré (morphème libre *si* en français, morphème lié du comparatif -*er* en anglais…), serait la plus caractéristique car elle établit un lien entre la syntaxe et la sémantique. Dans certaines langues telles le japonais ou l'irlandais, les adjectifs possèdent d'ailleurs d'autres caractéristiques syntaxiques propres, qui les distinguent des autres catégories lexicales (Matushansky 2005 : 34).

On sait toutefois qu'il existe des langues sans adjectifs, telles le nez-percé, langue amérindienne (Coppolani 2018 : section 2.3.2.1). D'autres présentent des adjectifs ne satisfaisant pas tous les critères : par exemple en slave ou slavé (langue amérindienne du groupe athapascan/déné, n'ayant

[1] Le caractère facultatif inclus dans une telle définition sert, selon l'auteur, à exclure les déterminants, qui modifient bel et bien le nom, mais qui servent essentiellement à introduire le syntagme nominal. Les déterminants sont en principe obligatoires, bien qu'ils puissent être omis dans le cas de la plupart des noms propres, des autonymes, et des noms communs dans certains contextes, comme les constructions à verbe support (Gaatone 2016 : 17).

aucun rapport avec la branche slave des langues indo-européennes), les adjectifs ne satisfont pas le premier critère, en russe ils ne satisfont que le deuxième (Matushansky 2005 : 33). Dans la mesure où les critères de définition diffèrent d'une langue à l'autre (selon que l'on considère ses propriétés comme principalement sémantiques, morphologiques ou syntaxiques), on ne peut donc affirmer que la catégorie de l'adjectif soit universelle.

2.2. Aspects syntaxiques

2.2.1. Fonctions

L'adjectif dépend d'un autre terme de la phrase, nom ou pronom. La fonction qu'il revêt se définit selon la relation qu'il entretient avec ce terme. Elle peut être de quatre sortes (Haßler 2016 : 27-29).

- **attribut** ➔ **du sujet** (facultatif ou non) : qualifie le sujet par l'intermédiaire d'un verbe dit d'état ou attributif. La relation attributive désigne une caractéristique inhérente (*Elle est intelligente. Elle se fait vieille*) ou accidentelle du sujet (*Elle est rentrée fatiguée*).
  ➔ **de l'objet** (facultatif ou non) : qualifie l'objet par le biais d'un verbe transitif.
- **modifieur** ➔ **épithète** (antéposé ou postposé au nom) : modifieur facultatif dans le syntagme nominal. On trouve également le terme d'épithète liée, adjointe au nom sans pause orale ni écrite, solidaire du nom et donc non autonome. Elle est facultative sur le plan syntaxique (la suppression est possible) mais pas toujours sur le plan sémantique (la suppression de l'adjectif à valeur restrictive modifie le sens du référent).
  ➔ **apposé** (ou apposition). On trouve également le terme d'épithète détachée, séparée du nom par une pause à l'oral et par des virgules à l'écrit. L'apposition est mobile dans la phrase et peut notamment porter sur un pronom : *Toute surprise, elle n'en croyait pas ses yeux.*

Kyriakaki (2011 : 94) nomme le modifieur *épithète* modifieur **restrictif** et le modifieur *apposé* modifieur **non-restrictif / appositif**. Matushansky (2005 : 10-11) nomme la fonction attribut *prédicative*. Certains adjectifs peuvent ne revêtir que l'une ou l'autre des fonctions (seulement épithète : *nucléaire*). Matushansky (2005) explique que dans beaucoup de langues du monde, les adjectifs ont la même forme, qu'ils prennent l'une ou l'autre des fonctions. Dans d'autres, ils présentent une forme différente du point de vue morphologique : « l'accord peut être marqué sur les adjectifs épithètes sans être marqué sur les adjectifs prédicatifs comme en néerlandais ou en allemand […] ; ou les adjectifs épithètes peuvent porter une marque morphologique supplémentaire qui semble indépendante de l'accord, par exemple en russe » (Matushansky 2005 : 35).

On entend par **modifieur** tout élément adnominal qui porte sur la tête et fait partie du même syntagme nominal, apportant une information supplémentaire sur le référent mais pouvant souvent être supprimé. C'est pour cette raison que l'on dit qu'il *modifie* le nom. Ce peut être un adjectif ou

syntagme adjectival[1], un syntagme génitival, un syntagme prépositionnel, une proposition relative et une apposition nominale non séparée de la tête nominale par une virgule à l'écrit, une pause à l'oral ou une rupture intonationnelle. Il modifie soit le nom seul, soit le nom et les modifieurs qui sont proches de lui, en fonction de l'incidence de sa portée DEF [A [A N]], DEF [A [N A]], DEF [A [N SGEN]].

2.2.2. Distribution

Les questions les plus traitées en syntaxe sont la question de la position anté- ou postnominale de l'adjectif épithète et l'ordre des adjectifs épithètes lorsqu'ils sont plusieurs. En anglais et dans les langues germaniques, la position la plus fréquente de l'adjectif est anténominale. Dans certaines langues, la position la plus fréquente de l'adjectif est postnominale, ainsi dans les langues romanes et en hébreu (Matushansky 2005 : 31-32). On y trouve le même ordre des mots… mais inversé.

De nombreuses études ont porté sur la position de l'adjectif en français, qui modifie son interprétation (entre autres Forsgren 1978, Wilmet 1981a, Wilmet 1981b, Forsgren 1997, Bouchard 1998, Abeillé et Godard 1999, Laenzlinger 2000 et Laenzlinger 2005, Nølke 2001, Thuilier 2012 et Thuilier 2015). On peut lire souvent que l'adjectif prend son sens propre en position antéposée et postposée, son sens figuré en position antéposée. Mais ce n'est pas le cas de tous les adjectifs et de tous les syntagmes, ce qui laisse croire à une absence de différence de sens. (Nølke 2001 : 168-169) résume l'interprétation communément admise de l'adjectif épithète en français en fonction de sa position.[2]

- L'antéposition correspond à une valeur subjective, expressive, banale, affaiblie, figurée, numérale, déterminative de l'adjectif, voire à une fusion de l'adjectif avec le nom pour former une seule pensée (mot composé ou quasi mot composé), exprimer une propriété inhérente du nom.
- La postposition correspond à une valeur objective, distinctive, qualificative, classificatrice, pleine, littérale, véhiculant une information nouvelle, exprimant une pensée distincte de l'idée exprimée par le substantif, éventuellement remplaçable par une subordonnée relative.

La « nature » de l'adjectif jouerait également un rôle : seraient antéposés les adjectifs dits élémentaires, ayant un sens large ou vague, à usage fréquent. Seraient postposés les adjectifs relationnels, de couleur, ayant un sens précis ou technique (Nølke 2001 : 169). Le syntagme nominal joue lui-même un rôle : l'antéposition est plus fréquente si le syntagme est défini, si le syntagme est sujet de la phrase, si le nom est composé ou polysyllabique. La postposition est plus fréquente si le syntagme est indéfini, s'il est complément verbal ou phrasal, si le nom est simple ou monosyllabique, si l'adjectif régit son propre complément (Nølke 2001 : 169). En résumé, il faut compter avec les facteurs suivants : le sémantisme de l'adjectif, celui du nom, la structure morpho-syntaxique de l'adjectif, celle du nom, la position et la fonction syntaxiques du syntagme nominal dans la phrase… mais aussi la rection syntaxique interne au syntagme nominal et la prosodie (Nølke 2001 : 176).

[1] Ce qui inclut un participe converti en adjectif ou syntagme participial.

[2] Pour un état de l'art de différentes théories grammaticales, structuralistes, fonctionnalistes et cognitivistes sur la place de l'adjectif épithète en français, voir Nølke (2001 : 168-175).

Les analyses de Nølke (2001) rejoignent celles de Cinque (2010), qui les développe dans un cadre théorique très différent. Son livre analyse en détail la place de l'adjectif en anglais (et autres langues germaniques) et en italien (et autres langues romanes). Il observe les mêmes effets de sens lorsque l'adjectif est antéposé en anglais et lorsqu'il est postposé en italien ; lorsque l'adjectif est postposé en anglais et antéposé en italien. Les critères d'analyse sémantique sont les suivants :

- interprétation en un sens général (*individual-level*) ou en un sens ponctuel (*stage-level*)
- interprétation restrictive ou non-restrictive
- interprétation modale ou comme une proposition subordonnée relative implicite
- interprétation intersective ou non-intersective (adverbiale)
- interprétation relative (à une classe de comparaison) ou absolue
- interprétation comparative ou absolue des superlatifs
- interprétation spécifique ou non-spécifique
- interprétation évaluative ou épistémique de l'adjectif *inconnu*
- interprétation dépendant (du syntagme nominal) ou anaphorique (du discours) de *différent*

L'auteur, grâce à sa liste de critères sémantiques, découvre une double opposition entre deux types de positions et deux types de langues. La structure du syntagme nominal y est organisée en miroir (Cinque 2010 : 2-3, 5-17, 91) :

Tableau 8 : Résumé de la comparaison entre les propriétés interprétatives de l'adjectif anténominal et postnominal en anglais et italien (Cinque 2010 : 16-17)

English (Germanic)		
Prenominal adjectives	N	Postnominal adjectives
stage-level or individual-level reading		stage-level (or individual-level) reading
restrictive or nonrestrictive reading		restrictive reading
implicit relative clause or modal reading		implicit relative clause reading
intersective or nonintersective reading		intersective reading
relative or absolute reading		[cannot be tested]
comparative or absolute reading of superlatives		[cannot be tested]
specificity- or non-specificity-inducing reading		specificity- or non-specificity-inducing reading
evaluative or epistemic reading of 'unknown'		[cannot be tested]
NP-dependent or discourse anaphoric reading of 'different'		[cannot be tested]

Italian (Romance)		
Prenominal adjectives	N	Postnominal adjectives
individual-level reading		individual-level or stage-level reading
nonrestrictive reading		restrictive or nonrestrictive reading
modal reading		modal or implicit relative clause reading
nonintersective reading		intersective or nonintersective reading
absolute reading		relative or absolute reading
absolute reading of superlatives		comparative or absolute reading of superlatives
specificity-inducing reading		specificity- or non-specificity-inducing reading

evaluative reading of ‘unknown’	evaluative or epistemic reading of ‘unknown’
NP-dependent reading of ‘different’	NP-dependent or discourse anaphoric reading of ‘different’

En anglais, la position anténominale est systématiquement ambiguë entre deux valeurs possibles pour chaque propriété sémantique, tandis que la position postnominale (lorsqu’elle est grammaticalement possible) n’offre pas d’alternative : une seule valeur est possible. C’est le contraire en italien et dans les langues romanes, où la position postnominale est systématiquement ambiguë, tandis que la position anténominale n’a qu’une valeur possible (Cinque 2010 : 17).

L’auteur postule un fonctionnement en réalité unique dans toutes les langues, comprenant deux sources syntaxiques : la structure sous-jacente donnant lieu à un adjectif épithète en surface peut être de deux types. La première est la position de spécifeur de projection fonctionnelle (*direct phrasal specifier of a dedicated functional head of the extended projection of the Noun*), qui correspond à la modification directe (nous verrons cette notion plus avant, en section 3.5 p. 97). La seconde est la position de subordonnée relative réduite (*predicate of a reduced relative clause*), qui correspond à la modification indirecte. Le Tableau 9 résume leurs propriétés interprétatives et syntaxiques (Cinque 2010 : 27-28, 33-34) :

Tableau 9 : Résumé des propriétés sémantiques et syntaxiques des deux sources de l’adjectif (Cinque 2010 : 27, 33)

	Indirect (reduced RC) modification	*Direct modification*	
[Det	[stage-level (or individual-level)	[individual-level	NP]]]
[Det	[restrictive	[nonrestrictive	NP]]]
[Det	[implicit reative clause	[modal	NP]]]
[Det	[intersective	[nonintersective	NP]]]
[Det	[relative (to a comparison class)	[absolute	NP]]]
[Det	[comparative (with superlatives)	[absolute (with superlatives)	NP]]]
[Det	[specificity- or non-specificity-inducing	[specificity-inducing	NP]]]
[Det	[epistemic ‘unknown’	[evaluative ‘unknown’	NP]]]
[Det	[discourse anaphoric ‘different’	[NP dependent ‘different’	NP]]]
[Det	[deictic	[generic	NP]]]
[Det	[literal interpretation	[possible idiomatic interpretation	NP]]]
further away from the N		closer to the N	
not rigidly ordered		rigidly ordered	
possible in predicate position		not possible in predicate position	

Certains adjectifs ne peuvent provenir que de l’une des deux sources et connaître donc une seule lecture possible. D’autres au contraire, notamment dans les langues romanes, peuvent provenir des deux sources différentes, selon le sens que le locuteur souhaite leur voir attribuer, d’où des ambiguïtés avec *pauvre, ancien, vieux, bon* (Cinque 2010 : 92-93). Cela explique les ambiguïtés de sens mentionnées en début de section. Les adjectifs de couleur constituent un cas particulier car leurs propriétés sémantiques et syntaxiques sont complexes (Cinque 2010 : 93). Employés au sens propre dans l’usage littéraire ou soutenu, ils sont des modifieurs directs anténominaux, non-restrictifs, non-

intersectifs (*les vertes prairies d'Irlande*). Employés au sens propre dans l'usage courant, ils sont des modifieurs indirect postnominaux, restrictifs, intersectifs (*les feutres verts*). Employés au sens figuré, ils sont des modifieurs directs postnominaux classificateurs (*un député vert*).

2.2.3. Ordre des adjectifs

L'ordre des adjectifs épithètes dépend surtout de leur sémantique. Morzycki (2015 : 44) affirme que « *across languages, these tend to appear in that order when prenominal (that is, evaluative < size < shape < color); postnominally, they tend to occur either in the same order or in its mirror-image* ». Cinque (2010 : 38), sur la base de Greenberg (1963) et de Cinque (2005), propose une typologie de langues partageant le même ordre des adjectifs épithètes (Figure 9).

a. *English, Chinese…*
$A_{size} > A_{color} > A_{nationality} > N$

b. *0*
$* A_{nationality} > A_{color} > A_{size} > N$

c. *Welsh, Irish…*
$N > A_{size} > A_{color} > A_{nationality}$

d. *Indonesian, Yoruba…*
$N > A_{nationality} > A_{color} > A_{size}$

Figure 9 : Ordre d'apparition des adjectifs dans différentes langues selon (Cinque 2010 : 38)

Pour le grec, Stavrou (1999 : 202-203) compile différents travaux sur l'ordre des adjectifs, dont ceux de Sproat et Shih (1988), et propose la hiérarchie sémantique de la Figure 10 :

numéral > qualité (évaluatif, modal) > orienté vers le sujet (manière) > taille > forme > couleur, origine ou nationalité, matière (l'ordre de ces trois derniers pouvant varier)

Figure 10 : Ordre d'apparition des adjectifs selon Stavrou (1999 : 203)

La syntaxe est déterminée par le type d'adjectif, par la sémantique et par des considérations d'ordre pragmatique. Stavrou (1999 : 202-203) distingue les adjectifs orientés vers l'objet, c'est-à-dire délivrant une propriété de celui-ci, une caractéristique objective de la réalité, des adjectifs orientés vers le locuteur, c'est-à-dire délivrant une pensée, une considération subjective, un jugement de la personne qui parle. Les premiers sont les plus proches du nom que les seconds (Figure 11) :

$$\text{Det} > \text{Adj}^{\text{locuteur}} > \text{Adj}^{\text{objet}} > \text{Nom}$$

Figure 11 : Ordre des constituants nominaux

Plus un adjectif doit s'interpréter de manière absolue, plus on peut s'attendre à ce qu'il soit proche du nom qu'il modifie (Stavrou 1999 : 204). « *Thus, the AP adjacent to the head is the one that is intended to 'classify' the noun, while every preceding AP has scope over the following adjective+(adjective+)noun. Generally, the higher an AP appears, the less object-like the property it*

denotes » (Stavrou 1999 : 220). Le rôle de l'adjectif épithète est en effet d'aider à identifier le référent du nom. S'il y en a plusieurs, l'ordre dans lequel ils apparaissent est lié à leur potentiel d'identification de la référence : d'abord les adjectifs ayant un potentiel moindre (orientés vers le locuteur), ensuite ceux ayant un potentiel important (orientés vers l'objet). Cet ordre peut sembler étonnant en ce qu'il retarde l'identification du référent. À la différence de sens des deux types d'adjectifs correspond en réalité une hiérarchie dont le centre de gravité est le nom-tête. C'est que les adjectifs orientés vers l'objet visent à sélectionner le référent exact du discours parmi d'autres référents similaires (qui partagent la caractéristique du nom mais non la propriété de l'adjectif). En cela, ils sont plus importants que ceux orientés vers le locuteur parce qu'ils remplissent la fonction initiale et principale de l'adjectif, délimiter la référence du nom-tête. Ils s'accolent donc directement à lui. Certains peuvent en outre revêtir une fonction classificatoire, par exemple *vin rouge ~ vin blanc ~ vin rosé* – classification que n'opère pas, par exemple, *vin bouchonné*. Il s'agit des adjectifs qui ressemblent à des classifieurs (*américain, politique*) et des adjectifs qui dénotent l'origine, la matière, la forme ou la couleur d'objets (*métallique, gros, bleu*). Ils précèdent alors directement le nom et forment un concept unitaire, un lexème unique avec lui (Stavrou 1999 : 206-207), à l'instar d'un mot composé : *vin rouge* fonctionnerait comme *pause-déjeuner* et *Alexandre le Grand* comme *Alexandre Dumas*.

Cet ordre canonique peut être violé en cas de postposition (Stavrou 1999 : 215-216) :

(1) Grec (Stavrou 1999 : 215-216)

ena	*oreo*	*forema*	*metaksoto*
a	nice	dress	silk
ena	*metaksoto*	*forema*	*oreo*
a	silk	dress	nice

L'adjectif orienté vers l'objet (*metaksoto*, en soie) précède en (1) celui orienté vers le locuteur (*oreo*, beau). Les deux types d'adjectifs peuvent se trouver postposés au nom (Stavrou 1999 : 217-219). L'exemple (2)a comprend deux adjectifs orientés vers l'objet et qui sont postposés, tandis que l'exemple (2)b comprend un adjectif orienté vers le locuteur et qui est postposé. Dans la pratique, on postpose rarement plus de deux adjectifs orientés vers l'objet (Stavrou 1999 : 218).

(2) Grec (Stavrou 1999 : 217, 219)

a.	*sedonja*	*polihroma*	*mona*
	sheets	multi-coloured	single
b.	*ena*	*vivlio*	*endhiaferon*
	a	book	interesting

Un syntagme comme (2)b pose problème à l'autrice, car sa théorie prédit que « *Ideally, the adjective that follows the (raised) noun should always be of the object-oriented type* ».

Pour l'anglais, Laenzlinger (2000 : 59-60), s'appuyant partiellement sur Scott (1998), propose l'ordre en (Figure 12) pour les différents types d'adjectifs.

a.	Ordre des classes d'adjectifs	$Adj_{speaker\text{-}oriented}$ > $Adj_{subject\text{-}oriented}$ > Adj_{manner} > $Adj_{thematic}$
b.	Ordre des adjectifs dénotant les noms	$Adj_{quantification}$ > $Adj_{quality}$ > Adj_{size} > Adj_{shape} >Adj_{color} > $Adj_{nationality}$
c.	Ordre des classes d'adjectifs regroupés par méta-classes sémantiques	[$_{QUANTIF}$ Ordinal > Cardinal] > [$_{SPEAK\text{-}ORIENT}$ Subjective Comment > Evidential] > [$_{INTERNAL\ PHYSIC\ PROPERTY}$ Size > Length > Height > Speed > Depth > Width] > [$_{MEASURE}$ Weight > Temperature > ?Wetness > Age] > [$_{EXTERNAL\ PHYSIC\ PROPERTY}$ Shape > Color > Nationality/Origin > Material]
d.	Ordre des constituants du SN	D > Adj_{quant} > Adj_{qual} > Adj_{size} > Adj_{form} > Adj_{color} > $Adj_{nationality}$ > N

Figure 12 : Ordre d'apparition des adjectifs en anglais (Laenzlinger 2000 : 59-60)

Pour le français, Nølke (2001) propose une analyse des suites d'adjectifs : antéposés ou postposés, ils peuvent avoir deux relations structurales au nom, la *subordination* et la *coordination*.[1]

Avec adjectif antéposé :

> « Dans le cas de subordination, les adjectifs extérieurs ne modifient pas directement le substantif : ils concernent le groupe entier constitué par le substantif et les adjectifs intérieurs. Dans le cas de coordination, par contre, chaque adjectif caractérise le substantif directement et indépendamment des autres adjectifs […] :
>
> un bon vieux petit homme
> cette étroite, montueuse, raboteuse, sale et vénérable rue saint Jaques [sic]
>
> Les deux exemples font preuve des structures suivantes :
>
> (1') (A (A (A (S))))
> (2') ((A, A, A, A et A) (S)) » (Nølke 2001 : 199)

Avec adjectif postposé :

> « Comparons les deux exemples sous (3) et (4) :
>
> une solution commune européenne
> un morceau ennuyeux et triste et douloureux de ma vie passée
>
> Dans (3), il s'agit d'une solution qui est commune et d'une solution commune qui est une solution commune européenne ; ces deux qualités forment un tout solide et ordonné. […] Dans (4), il s'agit d'un morceau qui est successivement considéré comme un morceau ennuyeux ; un morceau triste et un morceau douloureux, sans que ces qualités soient intimement réunies. Les deux exemples font preuve des structures suivantes :
>
> (3') (((S) A) A)
> (4') ((S) (A et A et A)) » (Nølke 2001 : 209)

Dans le cas de la subordination, tous les adjectifs appartiennent au même intonème ; celui dont l'information est la plus importante est placé le plus près du substantif. Dans le cas de la coordination, les adjectifs sont « reliés » entre eux par une pause (coordination *asyndétique*), transcrite à l'écrit par

[1] On peut également parler de relation *hypotaxique* et *parataxique*.

une virgule, ou bien reliés par une conjonction de coordination (coordination *analytique*) (Nølke 2001 : 199, 209-210).

En **antéposition**, la subordination est plus fréquente que la coordination (Nølke 2001 : 200, 202).

- **Subordination** : Pour des raisons cognitives, on ne peut trouver plus de trois adjectifs subordonnés : plus d'adjectifs il y a, plus difficile sera l'interprétation du syntagme nominal (Nølke 2001 : 200). On trouve en général les adjectifs « élémentaires » au plus près du substantif. Un changement de l'ordre des adjectifs entraîne une modification de sens, si minime soit-elle : *une bonne vieille veste ≠ une vieille bonne veste*. Le premier syntagme a une connotation positive que n'a pas le deuxième (où certes la veste est bonne, mais elle est vieille). En effet, l'adjectif subordonné le plus proche du nom fusionne avec lui, « de sorte que le groupe parvient à dénoter un seul concept (nouveau) » (Nølke 2001 : 201-202). « Le mécanisme de subordination traite en bloc les adjectifs intérieurs et le substantif » (Nølke 2001 : 202). De plus, la structuration du domaine de focalisation prévoit déjà cet effet de fusionnement pour l'adjectif antéposé (cf. section 3.5, p. 97), donnant lieu à des figements tels que *jeune homme, jeune femme*, voire à des mots composés tels que *grand-mère, gentilhomme, rouge-gorge* (Nølke 2001 : 202).
- **Coordination** : La coordination est relativement rare, notamment celle **asyndétique** (Nølke 2001 : 202, 204) car « la petite pause (reflétée par une virgule dans la langue écrite) pose des problèmes à la perception de la pré-zone comme constituant un seul intonème » (Nølke 2001 : 202). Elle doit comprendre des adjectifs dits élémentaires : *mes <u>profondes, longues, insensibles</u> variations*. Grâce à la virgule, on peut changer l'ordre des adjectifs sans changement notable de sens, mais on peut aussi, en théorie, les accumuler sans restriction.
  Bien plus fréquente est la coordination **analytique** en *et* (*la <u>sinistre et atroce</u> surprise*), *ou* (*une <u>forte ou assez forte</u> avance*), plus rarement en *mais* (*son <u>vieillissant mais infatigable</u> époux*), qui introduit une opposition se heurtant à la tendance de la pré-zone « à présenter une seule pensée » (Nølke 2001 : 204). La coordination analytique en antéposition peut conduire à des expressions toutes faites (*un <u>seul et même</u> amour*). Il arrive que se combinent les coordinations asyndétique et analytique (*<u>bon, grand et noble</u> cœur*).

- Dans de rares cas, subordination et coordination se **combinent** (*la <u>fraîche et belle jeune</u> femme*). La seule structure possible est alors ((A et A) (A (S))) où la coordination est subordonnée à la subordination (Nølke 2001 : 206). Toute autre structure est exclue, car « la subordination implique toujours le substantif alors que la coordination forme un constituant indépendant » (Nølke 2001 : 206). On constate que pour des raisons cognitives, « c'est l'élément subordonné extérieur qui se compose d'une coordination », et que la cohésion entre l'adjectif intérieur et le substantif est forte.

En **postposition**, la coordination est nettement plus fréquente que la subordination (Nølke 2001 : 212, 215). En effet, la structure plus lâche de la post-zone entrave la subordination, possible seulement avec une épithète modificatrice subordonnée à une épithète relationnelle, éventuellement avec deux épithètes modificatrices. Elle favorise donc la coordination entre plusieurs adjectifs (de même fonction) ou entre adjectifs et d'autres types de modifieurs (Nølke 2001 : 212).

- **Subordination** : Elle se caractérise par sa forte cohésion, son appartenance au même intonène et au même domaine de focalisation. Elle se heurte à l'arborescence de la structure de focalisation et d'accentuation, ce pourquoi les exemples sont très rares.

  Subordination : (((S) A) A)
  Domaine de focalisation : (S (A (A)))

  Elle ne permet pas plus de deux adjectifs, dont le premier doit être restrictif, ou bien dont l'un des deux doit être relationnel et contigu au substantif : *réponse positive rapide, aventure galante illicite* ; *expression géographique abstraite, corps préfectoral actuel* (Nølke 2001 : 213, 214). Ainsi dans le syntagme *une expédition militaire désastreuse*, l'adjectif militaire « entre dans une relation si étroite avec son substantif qu'il forme avec celui-ci, pour ainsi dire, un nouveau substantif que le deuxième adjectif pourra qualifier » (Nølke 2001 : 214). L'auteur parle de groupe « subcatégorisé », où l'adjectif restrictif ou relationnel acquiert une propriété de « subcatégorisation » si forte qu'il permet l'adjonction d'un troisième adjectif : *l'art grec classique tant adoré* (Nølke 2001 : 214). Dans une suite d'adjectifs postposés, le dernier a une relation beaucoup plus lâche au nom que les autres. On peut ainsi poser un principe de « précision croissante », où le deuxième adjectif revêt un sens plus précis que le premier et où « l'adjectif le plus bas sera toujours focalisé parce qu'il constitue le dernier élément de la chaîne », ainsi que l'exige la structuration du domaine de focalisation (Nølke 2001 : 215).

- **Coordination** : La coordination permet aux éléments de se comporter comme un seul constituant syntaxique, parfois doté de son intonème propre (Nølke 2001 : 216). On trouve aussi bien de la coordination **analytique** (*un homme petit et vigoureux, l'état liquide ou gazeux, une voix jolie mais trop menue*) que la coordination **asyndétique** à l'aide d'une « virgule-c »[1] et d'une montée mélodique particulière (*un chemin montant, sablonneux, malaisé*, *un livre beau, grand, enrichissant, consolant*) (Nølke 2001 : 216). On peut voir s'accumuler les adjectifs juxtaposés à l'aide de la virgule-c[1] qui produit un effet d'énumération (Nølke 2001 : 218). C'est plus rarement le cas de l'accumulation d'adjectifs coordonnés : *un morceau ennuyeux et triste et douloureux de ma vie passée*. Plusieurs raisons à cela : la lourdeur stylistique de la répétition, mais aussi la cohésion forte et la relation binaire qu'engendre la coordination et qui rend difficile d'élargir le syntagme (Nølke 2001 : 217, 218). C'est pourquoi on peut dire que la coordination asyndétique est « ouverte » et se prête à un enchaînement virtuel, tandis que la coordination analytique est « fermée » et indique la clôture de la construction (Nølke 2001 : 218). « L'épithète relationnelle forme avec le substantif une unité binaire fortement cohésive qui dénote un seul (nouveau) concept » (Nølke 2001 : 220), les deux adjectifs coordonnés doivent être sémantiquement compatibles et de même type, quel que soit leur ordre. Ils peuvent être deux épithètes modificatrices, deux épithètes relationnelles (c'est rare mais possible : *la presse française et étrangère*), jamais les deux ensemble (**une crise ministérielle et grave*, **un parc municipal et immense*) (Nølke 2001 : 220). Si les adjectifs ne sont pas compatibles, le nom se prête alors à une lecture distributionnelle. Celle-ci est inhérente aux syntagmes au pluriel – avec ou sans coordination : « *une foule de monstres sociaux, juridiques, pécuniaires, religieux, gynécologiques,*

[1] Pour la distinction entre virgule-c et virgule-r, voir section 2.3.4.1, p. 171.

terriblement conséquents avec eux-mêmes, et d'ailleurs fort bien liés entre eux. » (Nølke 2001 : 221).
Il est fréquent que se combinent les coordinations asyndétique et analytique, avec une virgule entre les premiers adjectifs et une coordination entre les deux derniers (*un jeune homme blond, grand et mince*) (Nølke 2001 : 216). On peut parler de coordination trinaire. La coordination permet de combiner plusieurs types de modifieurs, comme l'adjectif et la proposition relative : « *une femme jeune encore et qui disait avoir vingt-huit ans* » (Nølke 2001 : 221).

- Il est possible que se **combinent** subordination et coordination : *une famille suisse excellente et charmante, sa figure pâle toujours aussi noble et aussi mélancolique* (Nølke 2001 : 221, 222). Dans ce cas, le premier élément doit subcatégoriser et le deuxième être plus précis que le premier ; le premier élément est donc la subordination et le deuxième la coordination. Le contraire est très rare : *une base populaire et sociale puissante*, au point que l'on préfère en général antéposer la subordination : *une puissante base populaire et sociale* (Nølke 2001 : 222, 223).

2.3. Aspects sémantiques

De nombreux classements des adjectifs ont été dressés, qui se recoupent souvent mais peuvent aussi se contredire. La sémantique de l'adjectif combinée à celle du nom représente un champ d'études ancien et très vaste, par exemple (Kamp 1975, Wilmet 1981a, McNally et Kennedy 2008).

2.3.1. Classes d'adjectifs

1) Nous avons vu que la classification de Stavrou (1999 : 202-203) distingue d'une part les **adjectifs orientés vers l'objet/le référent**, comprenant les qualificatifs (taille, forme, couleur…), les classificatoires (qui peuvent également dénoter la taille ou la couleur), les relationnels ou ethniques (cf. section 2.4.3, p. 185), d'autre part les **adjectifs orientés vers le locuteur** de l'énoncé. Les premiers ont un sens absolu, stable, car ils dénotent une propriété vérifiable de l'objet, les seconds un sens relatif. Elle propose l'expression grecque *eftelis politiki satira*, satire politique médiocre : le premier adjectif (*eftelis*, moyenne, médiocre) est orienté vers le locuteur car il représente son opinion et son jugement concernant le référent du nom. Le deuxième adjectif (*politiki*, politique) est orienté vers le référent car il dénote une propriété inhérente de celui-ci. L'adjectif orienté vers le référent, s'il est adjacent au nom et a une fonction classificatoire (*satire politique* est un type de satire), peut s'incorporer au nom pour former un quasi mot composé, *syntactic compounding* (Stavrou 1999 : 218).

Un tel regroupement permet une économie des distinctions sémantiques traditionnelles : adjectifs de taille, de couleur, de nationalité… Il permet d'expliquer des faits de de morphologie (les adjectifs orientés vers le locuteur acceptent mieux le changement de degré que les autres), de morpho-syntaxe (les adjectifs orientés vers le locuteur acceptent mieux la modification par un adverbe) et de syntaxe (ordre interchangeable de certains adjectifs et non d'autres, selon le type d'adjectif).

2) Nølke (2001 : 210-211) propose de distinguer les adjectifs des **pseudo-adjectifs**. Les premiers présentent une fonction modificatrice, sont graduables, antéposés ou postposés. Il s'agit en fait de la catégorie traditionnelle des **adjectifs qualificatifs**. Les seconds présentent une fonction relationnelle et ne sont pas graduables ni antéposables. Ils dérivent de noms et exercent la fonction sémantique de complément du nom (d'où « pseudo-adjectifs ») : *catholique = du catholicisme*. Ils dénotent par exemple l'appartenance à une ethnie ou la nationalité.

3) Thuilier (2012 : section 3.3.4) propose de distinguer des classes « lexicales » en français :

- **adjectifs désignant une propriété objective** (forme, couleur, nationalité, religion, catégorie sociale ou administrative…), qu'elle nomme également **adjectifs intersectifs** (*rond, fragile*) (Thuilier 2015). Ce sont les adjectifs qualificatifs, qui fournissent une qualité ou propriété intrinsèque, essentielle ou accidentelle, du référent.
- **adjectifs intensionnels**, qui ne dénotent pas une propriété mais « modifient la dénotation du nom ou la manière dont elle est vérifiée par le référent » (*vrai, supposé, futur, parfait*). Cette première distinction rappelle celle de Stavrou (1999) entre adjectifs orientés vers le référent et adjectifs orientés vers le locuteur de l'énoncé, mais ne la recoupe pas totalement. Par exemple, dans le syntagme « *le futur président* », « l'adjectif futur ne porte pas sur l'individu dénoté par président, mais sur l'intervalle de temps à partir duquel l'assignation de la propriété président est adaptée pour l'individu désigné » (Thuilier 2012 : 115).
- **adjectifs subsectifs**, dénotant une propriété dont l'interprétation varie selon le nom (*petit*).
- **adjectifs évaluatifs**, longs, morphologiquement complexes et à la position variable en français (*extraordinaire*).

Thuilier (2012 : section 3.8) propose de distinguer également des classes « sémantiques » :

- **adjectifs qualificatifs**, qui expriment la mise en relation d'une entité avec une propriété, peuvent revêtir en français toutes les fonctions d'un adjectif (épithète, attribut du sujet, attribut de l'objet, apposé), peuvent être antéposés ou postposés au nom lorsqu'ils sont épithètes. Ils peuvent occuper toutes les fonctions et sont modifiables en degré.
- **adjectifs relationnels**, qui « se distinguent des adjectifs qualificatifs par leur comportement sémantique, syntaxique et morphologique » en français (Thuilier 2012 : 131). Ils ne peuvent être ni attributs, ni apposés, ni épithètes antéposées, ni modifiables en degré, ni coordonnés à un adjectif qualificatif. Ils expriment une relation entre deux notions, deux entités, qui sont deux noms de manière sous-jacente – c'est pourquoi ils sont l'équivalent d'un complément du nom. Un *animal préhistorique* est un *animal de la Préhistoire*, le *système solaire* est le *système du soleil*. Ils sont le plus souvent dérivés de noms, d'où leur longueur syllabique (*organisationnel, pétrolier, présidentiel*). Ces adjectifs se soudent avec le nom pour composer un nouveau terme, à la limite du nom composé (*fièvre aphteuse*). La grammaire traditionnelle les nomme également adjectifs classificatoires, classifiants ou objectifs, car ils expriment l'appartenance à une classe de façon objective et externe au locuteur.

  L'autrice critique ce type d'adjectifs : tous les dénominaux ne sont pas relationnels (*ce voyage original*) et certains relationnels ne sont pas dénominaux (*le logement étudiant*). Certains

relationnels peuvent revêtir la fonction attribut (*Cette revue est mensuelle*), être modifiés par un adverbe (*une revendication strictement syndicale*), être coordonnés à un qualificatif (*un véritable problème organisationnel et récurrent*), voire être antéposés au nom pour créer un effet stylistique (*avec un présidentiel mépris*). De tels adjectifs « ont perdu leur dimension relationnelle, c'est-à-dire qu'ils n'expriment plus une relation. Dans ce cas, on ne peut pas définir une classe lexicale de relationnels, mais on doit parler d'une classe d'emplois relationnels qui se caractérisent notamment par les propriétés syntaxiques et sémantiques définies plus haut » (Thuilier 2012 : 133).

- **adjectifs ordinaux** (*premier, second, deuxième, troisième*).
- **adjectifs indéfinis**, « catégorie intermédiaire entre les déterminants indéfinis et les adjectifs qualificatifs » (*différent, autre, certain, quelconque, divers, tel, même*) (Thuilier 2012 : 134).

4) Matushansky (2005 : 22), dans son état de l'art sur les types sémantiques d'adjectifs, présente les adjectifs **extensionnels**, dont « la sémantique de la modification peut toujours être exprimée par une opération sur l'ensemble dénoté par le nom (i.e. sur son extension) » (adjectifs intersectifs, subsectifs). Elle les distingue des adjectifs **intensionnels**, qui peuvent s'interpréter de plusieurs façons. Ainsi l'exemple (3) est-il ambigu :

(3) Français (Matushansky 2005 : 23)
Olga est une belle danseuse.
a. *une danseuse et belle en tant que danseuse*
b. *une danseuse et belle en tant qu'être humain*

Elle qualifie ces deux grands types de « basiques : extensionnels [dépendants du contexte (tels que *petit*) ou pas (par exemple, *rouge*)] et intensionnels (par exemple, *futur, présumé*) » (Matushansky 2005 : 25). Elle avertit que les adjectifs intensionnels sont traditionnellement désignés comme « **évaluatifs** », terme également employé pour des adjectifs non prédicatifs et non compatibles avec un morphème de degré.

2.3.2. Interprétation du syntagme nominal

Les adjectifs les plus courants dans les langues du monde appartiendraient aux quatre catégories sémantiques suivantes : intersectif, subsectif, non subsectif, privatif, avec les sens de « DIMENSION (*petit, grand, court…*), AGE (*jeune, nouveau, vieux…*), VALEUR (*bon, mauvais, parfait…*) ou COULEUR (*rouge, blanc, noir…*) » (Matushansky 2005 : 13). Il s'agit en réalité d'interprétations du syntagme nominal déterminées par l'adjectif. Morzycki (2015 : chap. 2), s'appuyant sur Partee (1995), dresse une typologie de l'interprétation des syntagmes nominaux avec modifieur adjectival.

- **lecture intersective** : il s'agit d'une relation symétrique entre l'adjectif et le nom. Le référent appartient à la fois à l'ensemble des entités dénotées par l'adjectif et à l'ensemble des entités dénotées par le nom. Dans la phrase *Floyd est un chirurgien canadien*, le référent, Floyd, appartient à l'ensemble des {chirugiens} et à l'ensemble des {entités canadiennes} (Morzycki

2015 : 14). On entend par là que sa dénotation doit être obtenue en croisant deux ensembles de référents. « Si *rouge* dénote un ensemble d'individus (qui possèdent la propriété de refléter les vagues électromagnétiques de longueur entre 650 et 750 nanomètres) et *planète* dénote un autre ensemble d'individus, alors le NP *planète rouge* dénote l'intersection de ces deux ensembles » (Matushansky 2005 : 11). Ces adjectifs sont les seuls qui ont le même sens en position épithète et en position attribut : *Camille est une linguiste suisse* = *Camille est suisse* et *Camille est une linguiste*. Il ne s'agit pas tant d'une qualité de l'adjectif ou du syntagme nominal que d'un type de lecture. Le référent se trouve à l'intersection de deux ensembles ; c'est là que s'opère la sélection. La plupart des adjectifs produisant une telle interprétation sont des adjectifs qualificatifs de taille, de couleur, de forme…

(4) (Lekakou et Szendrői 2007 : 146)

a. X is an Adjective Noun ⇒ X is Noun

b. X is a good pupil ⇒ X is a pupil

On procède de même pour calculer le sens global d'un syntagme comprenant plusieurs adjectifs épithètes, comme *linguistes théoriques contemporains*. Stavrou (1999 : 203) baptise cette interprétation **conjointe, plate ou parallèle** (Figure 13).

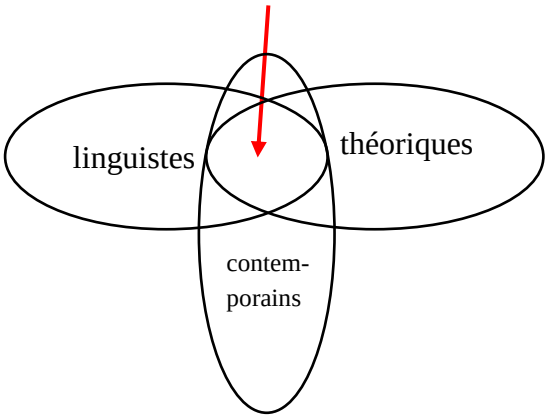


Figure 13 : Lecture intersective *alias* interprétation parallèle du syntagme nominal (Stavrou 1999 : 208)

Il existe des linguistes qui ne sont pas théoriques ni contemporains, ou qui sont soit l'un, soit l'autre, soit les deux. Il existe des entités contemporaines qui ne sont pas linguistes ni théoriques. Il existe enfin des entités théoriques qui ne sont pas linguistes ni contemporaines. *Linguistes théoriques contemporains* représente l'intersection des trois ensembles {linguistes}, {entités théoriques} et {entités contemporains}, c'est-à-dire l'intersection des référents du nom et de ses deux adjectifs. Cette interprétation se rencontre aussi dans les séries d'adjectifs juxtaposés par des virgules, où chaque adjectif modifie le nom indépendamment des autres. Leur ordre est assez libre et aisément modifiable ; leur nombre potentiel est très important, sans limite, car la modification « parallèle » est équivalente à la coordination (Stavrou 1999 : 210). Ces adjectifs sont orientés vers l'objet.

- **lecture subsective** : il s'agit d'une relation asymétrique entre l'adjectif et le nom car l'interprétation du premier dépend de celle du second. Dans la phrase *Floyd est un chirurgien habile*, le référent, Floyd, n'est pas au croisement de l'ensemble des {chirugiens} et de

l'ensemble des {entités habiles}. Il n'existe pas d'ensemble {entités habiles} indépendant, car *habiles* ne désigne pas la même réalité selon qu'il s'applique à un chirurgien ou à un incendiaire. *Habile* prend un sens différent en fonction du nom auquel il est associé car son référent se trouve en réalité au sein de l'ensemble de ce nom. Le référent, ici Floyd, appartient en fait à l'ensemble des {chirurgiens} et au sous-ensemble {chirurgiens habiles}. On parle d'interprétation subsective car l'adjectif représente un sous-ensemble (*subset*) du nom. Sont subsectifs des adjectifs comme *cher, minable, typique.* Ils n'ont pas non plus le même sens en fonction épithète et attribut : *Camille est une bonne linguiste ≠ Camille est bonne + Camille est une linguiste* (Matushansky 2005 : 12). Stavrou (1999 : 203, 208) baptise cette interprétation **hiérarchique** ou *scope reading* (Figure 14).

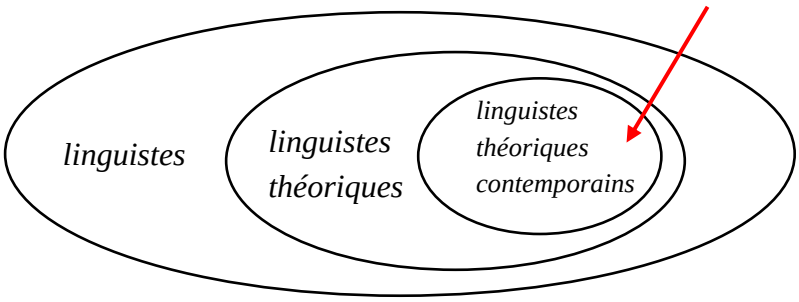


Figure 14 : Lecture intersective *alias* interprétation hiérarchique du syntagme nominal (Stavrou 1999 : 208)

Il n'existe pas de linguistes contemporains qui ne soient pas également théoriques – ni d'entités théoriques qui ne soient pas également linguistes. Chaque cercle désigne un ensemble de référents, où deux cercles sur trois sont imbriqués et forment des sous-ensembles (*subsets*) de l'ensemble supérieur. Lorsque le syntagme nominal donne lieu à une telle interprétation, l'autrice remarque qu'il comprend rarement plus de deux adjectifs orientés vers l'objet. Le nombre d'adjectifs successifs est selon Stavrou (1999 : 210) restreint, et leur ordre en est fixé par le sens du syntagme : les adjectifs orientés vers le locuteur précèdent ceux orientés vers l'objet, plus proches du nom sur l'axe syntagmatique. Ces adjectifs peuvent donc être orientés vers l'objet ou vers le locuteur.

- **lecture intersective en apparence subsective** : l'interprétation de certains adjectifs, notamment de taille, est à première vue subsective. Dans le syntagme *un petit éléphant*, la dénotation de *petit* semble dépendre de celle d'*éléphant* – puisqu'un petit éléphant est somme toute plus grand qu'une *grande souris*. Le sens de *petit* est à interpréter comme « petit-pour-un-éléphant », en relation étroite avec le sens du nom car la dimension de l'élément est relative. Mais les chercheurs s'accordent pour considérer ces adjectifs comme amenant une interprétation intersective du syntagme (Figure 15). Il semble qu'en employant les adjectifs de taille, on sous-entende toujours un standard de dimension approprié pour le nom. Ce standard est différent dans « grande-pour-une-gousse-d'ail » et « grande-pour-une-pomme-de-terre », ou dans « grande-pour-une-souris » et « grand-pour-un-éléphant ». Morzycki (2015 : 21) montre que ce standard ne dépend pas que du nom, et propose un test pour distinguer les adjectifs réellement subsectifs des adjectifs faussement subsectifs, à l'aide des prépositions anglaises *for* et *as*. Cinque (2010 : 18-19) considère également qu'il s'agit d'adjectifs intersectifs, même s'ils sont vagues et dépendants du nom et du contexte. La

dépendance du contexte vaudrait d'ailleurs aussi pour les vrais adjectifs subsectifs. Cinque (2010 : 19) propose même des exemples de contextes où des adjectifs intersectifs sont employés avec un usage en apparence subsectif : *Mary is blonde for a southern Italian (though not for a Scandinavian); That man is carnivorous for a vegan. (He eats fish)*. On voit que tout adjectif intersectif est relatif à un standard implicite et peut être employé à la manière d'un adjectif subsectif. Matushansky (2005 : section 3) approuve ces auteurs : « cette propriété [l'intersectivité] est masquée par le fait que [les adjectifs] sont vagues, c'est-à-dire que leur interprétation dépend du contexte. Ce type d'adjectifs est connu sous le nom d'adjectifs scalaires (*scalar, gradable*) ». Une telle dépendance contextuelle est traitée par les linguistes soit comme de l'imprécision (auquel cas ils ont le type sémantique des prédicats), soit comme une question de degré (auquel cas ils ont un type sémantique propre). Selon Matushansky (2005 : 23), il convient de « traiter une grande partie des adjectifs subsectifs comme des adjectifs intersectifs avec un argument de degré dépendant du contexte ». Ces adjectifs sont orientés vers l'objet.

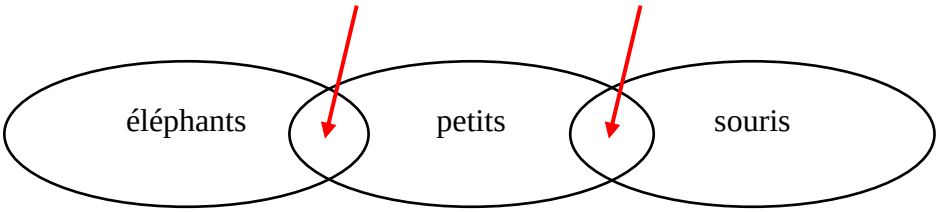


Figure 15 : Interprétations intersectives de l'adjectif *petit*

- **lecture ni intersective ni subsective** : Certains adjectifs amènent une lecture particulière car ils sont une « fonction » qui s'applique au sens du nom (Morzycki 2015 : 23) : *présumé, plausible, probable*... Ils n'ont pas le même sens en position épithète et en position attribut (ce qui n'est pas toujours possible) : *Camille est une future maman* ≠ **Camille est future* + *Camille est une maman* (Matushansky 2005 : 12). Cette dernière les nomme adjectifs non subsectifs simples. Notons que dans cet exemple, le syntagme ne dénote pas nécessairement un individu membre de l'ensemble des mamans : il peut en faire partie (si Camille a déjà eu un enfant) ou non (si son premier enfant n'est pas encore né). Cinque (2010 : 51) parle d'adjectifs modaux ou adverbiaux. « *Nonpredicative adjectives derive from adverbs (contained in relative clauses) rather than from adjectives in predicate position* […]. *Thus* the former president *could be thought of as deriving from something like 'the x (who was) formerly a president',* an alleged murderer *from something like 'an x (who is/was) allegedly a murderer', and so on* » (Cinque 2010 : 51). Ils sont orientés vers le locuteur.

- **lecture privative** : Certains adjectifs ont une lecture négative, comme *feint, artificiel, fictif.* Selon Morzycki (2015 : 25), ils n'ont pas de sens intrinsèque ; ils ont pour fonction de nier celui du nom. Ils n'ont pas le même sens en position épithète et en position attribut : *Ceci est un faux diamant* ≠ *Ceci est faux* + *Ceci est un diamant* (Matushansky 2005 : 12). Cette dernière les nomme adjectifs non subsectifs privatifs. Notons que dans cet exemple, le syntagme dénote nécessairement un individu qui n'est pas membre de l'ensemble des diamants : un faux diamant n'est pas un diamant (Matushansky 2005 : 13). Ces adjectifs ne sont ni intersectifs,

ni subsectifs, ni prédicatifs (Cinque 2010 : 22). Ils sont orientés vers le locuteur.

Selon Cinque (2010 : 24), la distinction sémantique et syntaxique fondamentale se réduirait à celle entre adjectifs prédicatifs et non prédicatifs. À l'échelle du groupe nominal, elle correspond à la distinction entre adjectifs provenant d'une proposition relative réduite sous-jacente et adjectifs en modification directe (cf. section 3.5, p. 100) (Figure 16).

APs in reduced RCs (only predicative)	***APs in direct modification (only non-predicative)***
*intersective (*bald, blonde*,..)*	*adverbial (*former, total, mere*,..)*
*'for a N' subsective (*tall, big*,..)*	*modal (*possible, alleged*,..)*
*'as a N' subsective (*skillful, beautiful*,..)*	*privative (*false, fake*,..)*
They modify the referent of the nominal phrase	*They modify the reference of the nominal phrase (Bolinger 1967)*
They modify the extension of the nominal phrase	*They modify the intension of the nominal phrase (Siegel 1976a)*
The meaning of the AP is not bound to, or relative to, the meaning of the common noun	*The meaning of the AP is bound to, or relative to, the meaning of the common noun (Siegel 1976a, b; 1979)*
They modify something that already designates an individual	*They modify something that is still a predicate*

Figure 16 : Résumé des propriétés du syntagme adjectival selon sa source syntaxique (Cinque 2010)

3. La structure de l'information

L'information est le contenu propositionnel et sémantique du message linguistique, ce qu'il nous dit du monde, la conception qu'a le locuteur de la réalité. « La notion d'information est ainsi reliée intimement à la présence du locuteur et de l'allocutaire dans le texte. Le but des actes assertifs, par exemple, peut être de changer la conception qu'a l'allocutaire du monde » (Nølke 2001 : 86). Chacun a son univers de croyances, et l'information sert à modifier cet univers. Dans le cas d'une interrogation par exemple, l'information sert à compléter l'univers de croyances du locuteur, ou encore à remettre en cause celui de l'interlocuteur.

Du point de vue linguistique, l'information est structurée d'une manière propre à chaque langue. Certains éléments sont, eux, universels : toute information est véhiculée au moyen d'une langue, d'un code linguistique. Toute phrase est énoncée dans tel contexte communicatif au sein de tel co-texte discursif, bref dans un cadre de la communication constitué de la situation, des connaissances communes aux protagonistes, des présupposés, des informations connues : tout ceci facilite la compréhension de l'information essentielle. On a pu parler d'arrière-plan (*versus* premier plan) communicatif (Nølke 2001 : 88-89), mais ces termes ne renvoient pas exactement à la même chose selon les linguistes et les philosophes qui ont étudié la question.

3.1. Structure et statut de l'information

On trouve dans les études une opposition binaire entre une information déjà connue et une information nouvelle, éventuellement mise en valeur : thème ~ rhème, topique ~ commentaire, présupposition ~ focus, donné ~ nouveau… (Lyons 1999 : 228).

- Le couple **topique ~ commentaire** (Reinhart 1982), ou encore **thème ~ rhème** (Firbas 1964), constitue la structure de l'information à proprement parler. Il oppose ce sur quoi porte le message (topique ou thème), ce dont il va être dit quelque chose, au message lui-même, au poids informatif de la phrase (commentaire ou rhème).
- Le couple **présupposition ~ focus** relève de la *structure de l'information*. Cette expression, due à Halliday (1967) signifie la présentation structurée de l'information dans une phrase, découpée en différentes unités informatives. Les phrases déclaratives (par exemple : *Elle vient de m'appeler*) répondent à une question sous-jacente (*Qu'est-ce qu'elle a fait Nathalie ?*). Ce sur quoi porte la question (*qu'est-ce qu'elle a fait ? → vient de m'appeler*) est le focus ; le reste (Nathalie, le fait que Nathalie existe et le fait qu'elle ait fait quelque chose) est présupposé. Le focus peut être exprimé par des moyens prosodique, morphologique, syntaxique, lexical, ou de plusieurs ordres combinés (Büring 2009). Il établit ainsi une typologie détaillée de la manière de marquer le focus dans les langues du monde.
- Le couple **donné ~ nouveau** constitue le *statut de l'information*, en lien avec la structure de l'information car le topique, point de départ du message, est généralement déjà présent dans la conscience de l'interlocuteur : soit parce qu'il a déjà été mentionné en discours (donné), soit parce qu'il fait partie de la connaissance du monde (accessible), soit parce qu'il est présent

dans la situation d'énonciation (accessible). Au contraire, le commentaire consiste souvent à fournir une information nouvelle sur le topique. Cependant les couples de la structure et du statut de l'information ne coïncident pas complètement, car le commentaire peut également être constitué d'information donnée.

Le choix d'un topique est une décision de l'énonciateur dans l'organisation de son discours. C'est pourquoi il convient de distinguer le **topique** de la notion de **connu / déjà mentionné** – même si cette dernière est bien sûr un facteur qui influence la décision (Creissels 2006 : 111). Au contraire, même si le locuteur choisit ce dont il parle et s'il veut référer à telle ou telle entité, il ne choisit pas le statut de l'information, qui est déterminé par la connaissance de l'interlocuteur, sa mémoire, son attention, la conversation qui s'est déroulée auparavant (Gundel et Fretheim 2006 : 193). C'est pourquoi nous choisissons de séparer ce qui relève de la *structure* de l'information et ce qui relève du *statut* de l'information. La structure de l'information concerne la manière dont est transmise l'information, dont elle est mise en valeur ou atténuée. Le statut de l'information est référentiel : il s'agit de savoir si un référent a déjà été mentionné ou non dans le discours, ou bien s'il est supposé faire partie de la connaissance du monde de l'interlocuteur. Nous nous fondons également sur la distinction établie par (Götze et al. 2007 : 150) :

> « *"Discourse referents" are meant to comprise entities of many different types, that is individuals, places, times, events and situations, and sometimes even propositions. All these can be picked up by anaphoric expressions. Their information status reflects their "retrievability", which is meant to be understood as the difficulty of accessing the antecedent referent: a referent mentioned in the last sentence is easily accessible or "given", whereas one that has to be inferred from world knowledge is only "accessible" to the degree that the inference relation is shared between speaker and hearer. A discourse referent which lacks an antecedent in the previous discourse, isn't part of the discourse situation, nor is accessible via some relational reasoning has to be assumed to be "new".* » (Götze et al. 2007 : 150)

Soit l'expression « J'ai vu une montre rouge et une montre verte. J'ai acheté la verte ». L'expression *la verte* est sous focus contrastif, toutefois le référent est donné, puisqu'il vient d'être mentionné. Il convient également de distinguer le niveau **référentiel** du niveau **lexical**. Un référent peut être nouveau mais le lexème donné : « *J'ai défait mon lit, mis les draps à la lessive et rangé l'oreiller. Et toi, tu as défait ton lit ?* ». Au contraire, un référent peut être donné mais le lexème nouveau : « *J'ai enfin trouvé le temps de réparer l'armoire, ça m'a pris toute l'après-midi et j'ai même dû la démonter. C'est un meuble vraiment très lourd !* ». Le statut référentiel de l'information implique une relation entre une expression linguistique et l'entité non-linguistique correspondante dans la tête des interlocuteurs, le discours ou le monde réel (Gundel et Fretheim 2006 : 176). On parle également de présupposition existentielle, de référentialité, de condition de familiarité, d'identifiabilité, etc.

Structure et statut de l'information sont étroitement liés à la syntaxe. On peut observer les tendances suivantes dans les langues du monde :

- le topique est souvent le sujet de la phrase et/ou souvent placé en tête de phrase :

> « *An early tradition, going back to the Prague School (Bolinger 1954; Halliday 1967; Firbas 1966, 1974; inter alia) had suggested that when word-order is used above and beyond the signaling of*

> *GRs, a universal tendency existed to place the topic or theme earlier in the clause and the focus or rheme later.* » (Givón 1984 : 277)
>
> « La place initiale étant souvent occupée par un élément déjà connu de l'interlocuteur, alors que ce qui suit constitue une information nouvelle, celle qui possède donc une plus grande importance communicative. Ainsi, la tendance générale pour tout énoncé est d'aller du connu vers le nouveau. » (Lebas-Fraczak 2015 : 34)

- le commentaire est souvent le prédicat ou l'objet de la phrase… et/ou souvent placé en tête de phrase (stratégie contraire à la précédente).
- Il existe un lien entre topicalité et statut donné de l'information : le topique est souvent « présupposé » ou « défini » (voire générique). On y trouvera donc souvent des marqueurs de définitude tels que déterminant possessif, démonstratif, générique, article défini (Gundel et Fretheim 2006 : 179). Cela suggère que le topique doive être familier et identifiable, au sens où l'interlocuteur doit déjà avoir dans sa mémoire une représentation du référent du topique. Il s'agit d'une condition nécessaire mais non suffisante du topique (Gundel et Fretheim 2006 : 180, 181). Un tel lien n'existe pas entre le focus et le statut donné : le syntagme du focus peut être défini ou indéfini.

3.2. Focus et topique

Le focus et le topique sont des phénomènes pragmatiques qui relèvent du domaine linguistique. Ils sont universels en ce que toutes les langues semblent avoir un moyen de les coder, quel que soit ce moyen (Gundel et Fretheim 2006 : 191). Il s'agit de catégories à la fois pragmatiques/sémantiques et prosodiques/syntaxiques, donc grammaticales : « *contrary to what is sometimes assumed, however, the fact that topic and focus have pragmatic effects does not in itself make them essentially pragmatic* » (Gundel et Fretheim 2006 : 191-192). Une question intéressante serait de démêler quelles propriétés du topique et du focus sont purement linguistiques (*grammar-driven*) et lesquelles sont « *derivable from more general pragmatic principles that govern language production and understanding* » (Gundel et Fretheim 2006 : 193). On trouve de nombreux termes pour opposer la paire thème-rhème dérivée des travaux d'Aristote (Nølke 2001 : 89) (Figure 17) :

Le premier segment de l'énoncé	Le reste
Ce dont on parle	Ce qu'on en dit
Topique	Propos
Présupposé	Posé
Segment non accentué	Segment accentué
Substrat (ou base)	Focus
Connu	Non connu
Lié contextuellement	Non lié contextuellement
Information antérieure	Information nouvelle
Information peu importante	Information très importante
Peu de Dynamisme Communicatif	Beaucoup de DC

Figure 17 : Termes exprimant l'opposition *topic*/*focus* (Nølke 2001 : 89)

Les définitions varient notamment en fonction de la perspective psycholinguistique, linguistique

(phonologique, syntaxique, sémantique), pragmatique ou logique qu'adoptent les auteurs. Nous ne présentons pas ici les cas particuliers, tels que les phrases à information nouvelle (*all-new sentences*), les phrases sans topique (*topicless sentences*) ou sans focus (*focusless sentences*).

Le terme de *focus* peut avoir une signification différente en fonction des auteurs et des cadres théoriques. Il est souvent conçu comme un segment souligné par le locuteur, véhiculant une information essentielle. À l'origine, ce terme signifie « concentration », notamment « concentration sur un point » dans le langage technique, comme en photographie (Nølke 2001 : 91). Il renvoie à :

- ce qui est dit de quelque chose, plus exactement **ce qui est dit du topique** chez Lambrecht (1994) ou Dik (1997)
- le **foyer** de la focalisation simple chez Nølke (2001 : 93, 178)
- **l'élément informatif** du point de vue communicationel, le **rhème** chez Firbas (1964)
- « un élément de la phrase […] présenté comme particulièrement **chargé d'une valeur informative** » (Creissels 2006 : 111), notamment dans le cas de la réponse à une question incluant pronom ou adverbe interrogatif, ainsi que dans le cas d'une rectification
- une information nouvelle ou nouvellement introduite dans le discours – ***new-information-focus*** chez Götze et al. (2007)
- une information qui s'oppose à une autre – ***contrastive focus*** chez Götze et al. (2007)
- un événement pragmatique, la mise en relief d'une information par divers moyens linguistiques et dans un but expressif ponctuel – **focalisation spécialisée** chez Nølke (2001 : 93,178)
- le **prédicat logique**, ce qui est prédiqué du topique Gundel et Fretheim (2006 : 176)
- le **rhème** chez Vallduví (1992)
- plusieurs de ces éléments en même temps

Il est souvent opposé au *topique*, qui renvoie, selon les auteurs, entre autres à :

- **la chose dont on parle** chez Lambrecht (1994) ou Dik (1997) – ***aboutness topic*** chez Götze et al. (2007)
- le **substrat** de la focalisation simple chez Nølke (2001 : 93,178)
- « **le point de départ psychologique** et/ou positionnel », le **thème** chez Halliday (1985), selon Prévost (2003 : 52)
- **l'élément peu informatif**, du point de vue communicationel, le **thème** chez Firbas (1964), selon Prévost (2003 : 52)
- « un élément de l'énoncé à partir duquel l'énonciateur développe un **commentaire** » (Creissels 2006 : 110)
- **l'information déjà connue**, non encore introduite en discours (« une telle approche correspond souvent à un amalgame entre statut informatif et accessibilité cognitive, celle-ci n'étant en fait qu'une caractérisation plus ou moins nécessaire du topique » selon (Prévost 2003 : 52)

- « le **cadre spatial, temporel, individuel**, à l'intérieur duquel prend place la prédication principale » (Prévost 2003 : 54) – ***background*** chez Chafe (1976), ***frame topic*** chez Götze et al. (2007)
- le **complément du focus**, ce sur quoi porte la phrase (Gundel et Fretheim 2006 : 176)
- le **thème** chez Vallduví (1992)
- plusieurs de ces éléments en même temps

Trois facteurs conditionnent le topique selon Creissels (2006) : les connaissances partagées par les interlocuteurs, le contexte et notamment les éléments déjà présentés et connus, « les propriétés de topicalité que possèdent intrinsèquement les noms », a fortiori les noms propres (un humain ou un agent s'utilisant plus naturellement comme topique qu'un non-humain ou un patient). C'est une tradition qualifiée par Verhoeven et Skopeteas (2015 : 1) de « *form-to-function association between focus and constituent structure* ». On voit que les deux rôles informationnels sont conçus en opposition et que la partition se fait de façon binaire. Nous suivons ce modèle binaire et oppositionnel, en considérant que l'information se structure ainsi via la syntaxe.

3.3. La focalisation

Lorsque des constituants sont placés en position de focus, on parle de **focalisation** ; lorsqu'ils sont placés en position de topique, on parle de **topicalisation**.[1] Plus exactement, focalisation et topicalisation consistent à signaler explicitement un constituant jouant le rôle discursif de focus et de topique – selon des procédés qui varient d'une langue à l'autre (Creissels 2006 : 110). Gardiner (1932 : 119) explique que « dans une phrase, certains mots sont plus importants que d'autres. […] [Le locuteur] sait quelles choses il désire mettre en relief et quelles choses ne font que contribuer à son but et, pour marquer ces différences, il utilise les ressources offertes par l'ordre des mots, les modulations de la voix, la vitesse d'élocution, etc ». La focalisation se caractérise par trois propriétés constitutives et indépendantes (Nølke 2001 : 177).

- **syntagmatique** : le focus, ou foyer, se manifeste sur une partie continue de la chaîne linéaire (mot, syntagme, constituant) qui a donc une certaine étendue.
- **paradigmatique** : le focus est le résultat d'un choix fait entre les éléments d'un paradigme.
- **intentionnelle** : le choix du focus a été effectué dans un certain but, nommée *visée*.

Pour Nølke (2001 : 91), « la focalisation linguistique sert fondamentalement à présenter l'élément focalisé comme le résultat d'un choix fait par le locuteur dans le cadre d'un paradigme constitué d'éléments semblables ». Par conséquent, les traits sémantiques qui se trouvent dans le foyer ou focus seront soulignés, tandis que ceux qui se trouvent hors du foyer seront atténués. Le lexème focalisé garde son sens plein, le lexème non focalisé perd de sa valeur sémantique (Nølke 2001 : 180). Toute

[1] Creissels (2006 : 109) note que l'on trouve également les termes de rhématisation et de thématisation pour ces deux concepts.

focalisation constitue un acte d'identification et sert à cela (Nølke 2001 : 93, 178). La sélection est donc la fonction primaire de la focalisation (Nølke 2001 : 91) :

> « Grâce à la nature paradigmatique de la langue, on peut préciser que la focalisation linguistique sert fondamentalement à présenter l'élément focalisé comme le résultat d'un choix fait par le locuteur dans le cadre d'un paradigme constitué d'éléments semblables. Qui plus est, ce choix est toujours fait dans un certain but. Une conséquence immédiate de ces considérations semble être la nécessité de concevoir le focus comme étant le résultat d'un acte de focalisation, qui a lieu au moment de l'énonciation, et dont le « décodage » constitue une partie nécessaire du processus d'interprétation. » (Nølke 2001 : 91)

La focalisation est toujours structurée par une arborescence à gauche : $(_f \ldots (_f \ x_3 (_f \ x_2 (_f \ x_1))))$. L'auteur distingue :

- la **focalisation simple**, « dont la visée est la simple identification de l'élément focalisé », est « contrainte par un *domaine de focalisation* qui correspond à un groupe rythmique » (Nølke 2001 : 178, 197-198). Elle se réalise dans la structure syntaxique de la phrase (Nølke 2001 : 178). Son lieu se nomme le *foyer*, tandis que le reste de la phrase ne subit pas de focalisation et se nomme le *substrat* (Nølke 2001 : 143).
- la **focalisation spécialisée**, « où d'autres visées (spécialisées) s'ajoutent à l'identification », « n'est marquée qu'accidentellement par la syntaxe mais elle est toujours assistée par l'accentuation d'insistance » (Nølke 2001 : 178, 197). Les autres visées sont le contraste, la correction, le commentaire métalinguistique, la polarisation[1] (Nølke 2001 : 98-99). Elle est plus libre et il est possible d'en avoir plusieurs dans un même énoncé.

Le lieu de la focalisation est nommé *domaine* de la focalisation (appelé dans d'autres travaux *portée*, ou *scope* du focus). « Toute structure de surface d'un énoncé contient exactement un domaine de focalisation simple, ce qui revient à dire que tout énoncé renferme un et un seul focus simple » (Nølke 2001 : 96, 197). Cela n'empêche pas de trouver d'autres types de focalisation, à des niveaux différents (Nølke 2001 : 180) :

- la **focalisation majeure**, au niveau de l'énoncé,
- la **focalisation mineure**, au niveau d'un seul groupe rythmique.

Autrement dit, l'auteur met en parallèle la structure informationnelle d'un énoncé et celle d'un groupe tel que le syntagme nominal. Dans la phrase, la position de focus requiert l'existence de sèmes spécifiques qui distinguent des éléments similaires et réduisent par là le nombre de référents possibles à identifier. La fonction de complément du verbe, notamment dans une structure négative, requiert l'existence de sèmes spécifiques afin de les mettre en relief, ce qui réduit le nombre de référents possibles et permet l'identification. En effet, ce sont eux qui servent à établir des distinctions entre des éléments similaires. Au contraire, la position de topique requiert moins de sèmes : elle n'a pas pour rôle principal l'identification du référent. La fonction de sujet atténue ces sèmes car elle échappe à la focalisation majeure. « La focalisation majeure favorise très fortement la focalisation (mineure) de l'adjectif et, partant, sa postposition » (Nølke 2001 : 190).

[1] dans les langues germaniques notamment. Elle est nommée *Verum-Fokus* en linguistique allemande (Nølke 2001 : 99).

Le focus se manifeste sous de nombreuses formes « correspondant aux différents choix d'expression qui permettent d'indiquer le statut communicatif des éléments au sein des énoncés » (Lebas-Fraczak 2015 : 31). Voici quelques moyens de l'exprimer (Lebas-Fraczak 2015 : partie II) :

- moyens **morphologiques** : emploi de focalisateurs, ou morphèmes de focalisation (liés ou libres : particule, négation, conjonction…), sans autre changement dans la construction de la phrase. Dans certaines langues, il existe un morphème spécial pour la focalisation. Dans d'autres, on utilise un morphème qui sert à autre chose à l'origine : un article, un adverbe…
- moyens **syntaxiques** : l'ordre des mots, la relation de dépendance entre éléments, le type de phrase, une construction spécifique
- moyens **lexicaux** : emploi de lexèmes de focalisation, « qui remplissent ce rôle d'une manière sémantiquement explicite » (adjectif *important*, adverbe *surtout*, mot-phrase *attention*…)
- moyens paraverbaux **prosodiques** : changement d'intonation (montante ou descendante), changement de débit (accélération ou ralentissement), changement d'intensité (plus forte ou plus faible), accentuation (d'une syllabe plus qu'une autre, allongement vocalique, syllabisation), pause (courte ou longue)
- moyens paraverbaux **(typo)graphiques** : variation de formes des caractères (gras, italique, majuscules, capitales), soulignement, couleurs, signes de ponctuation, marquage de l'organisation en parties (espacements, marques numériques)

Beaucoup de langues combinent plusieurs moyens, par exemple intonation et changement de l'ordre des constituants, ou bien intonation et adjonction d'une marque de focalisation.

3.4. Deux types de focus

Les études distinguent en général deux types de focus, qui peuvent d'ailleurs coïncider dans la même phrase (Gundel et Fretheim 2006 : 182).

> « La sélection d'un élément à l'intérieur d'une classe est une opération que le locuteur effectue en permanence au cours de l'activité verbale. La focalisation consiste à rendre patente – à attirer l'attention sur – une de ces opérations de sélection. Je distinguerai deux types de focalisation, en fonction de la nature de la relation qu'entretiennent entre eux les éléments de la classe sur laquelle se fait la sélection. Si la relation concerne des éléments qui se trouvent en compétition pour une opposition donnée dans la chaîne, je parlerai de focalisation paradigmatique. Si la relation concerne des éléments co-présents dans la chaîne, je parlerai de focalisation syntagmatique. » (Queixalós 2000 : 403)

Les recherches fournissent de nombreux termes pour exprimer cette opposition (Figure 18) :

- *contrastive focus* (Givón 1990 : 221)
- *contrastive focus* (Götze et al. 2007)
- *contrastive focus* (Gundel et Fretheim 2006 : 181)
- focus paradigmatique (Queixalós 2000 : 403)
- focalisation contrastive (Creissels 2006 : 120)
- focalisation de contraste (Nølke 2001 : 99)

- *focus of assertion* (Givón 1990 : 221)
- *new-information-focus* (Götze et al. 2007)
- *information focus* (Gundel et Fretheim 2006 : 181)
- focus syntagmatique (Queixalós 2000 : 403)
- focalisation simple/informationnelle (Creissels 2006 : 120)

Figure 18 : Termes exprimant l'opposition entre focus contrastif et focus informatif

La notion de focalisation est souvent mise en lien avec une interrogation sous-jacente : le focus correspond à « l'information manquante dans une interrogation partielle » ou, dans une phrase affirmative, à « l'information qui comble une lacune », le constituant qui apporte la réponse attendue (Creissels 2006 : 120). On parle alors de focalisation **simple / informationnelle**. Götze et al. (2007) distingue focus de *nouvelle information sollicitée* (si une question a été explicitement posée par l'interlocuteur) et focus de *nouvelle information non-sollicitée* (sans interrogation explicite préalable). Elle est à distinguer de la focalisation **contrastive**, qui correspond à plusieurs structures communicatives impliquant « un contraste en l'information mise en focus et une information alternative » (explicitement présentée ou seulement présupposée) (Creissels 2006 : 120). Il en existe plusieurs : focalisation contrastive de *rejet*, de *substitution*, d'*expansion*, de *restriction* et de *sélection*. Götze et al. (2007) distingue de leur côté focus contrastif de *remplacement*, de *sélection*, d'*analyse* ou de *partonomie*[1] (*partiality*), d'*implication* et de *confirmation* (*truth-value - verum*).

Ne nous méprenons pas : toute focalisation est, en un sens, contrastive. Toutefois, « la focalisation 'simple' met le focus en contraste avec toute autre information a priori susceptible de figurer à la même place, alors que la focalisation 'contrastive' met en contraste deux informations particulières. » (Creissels 2006 : 120). Évoquer une **alternative** serait la fonction primaire du focus selon Rooth (1985) et Chafe (1976), cités par Gundel et Fretheim (2006 : 181). On pourrait toutefois considérer que la fonction primaire du focus étant **d'attirer l'attention**, l'effet de sélection est une conséquence parmi d'autres. « Les deux [types de focalisation] bien entendu ont pour finalité pragmatique d'indiquer que l'élément sélectionné est porteur d'une information considérée comme **proéminente** » (Queixalós 2000 : 403). C'est aussi l'avis de Vallduví (1992), cité par Gundel et Fretheim (2006 : 181). En effet, la fonction de sélection nécessite la **proéminence** de l'information mais l'inverse n'est pas vrai. Comme l'écrit Givón (1990 : 221), « *every contrastive focus is also the focus of assertion, but not vice versa* ». Gundel et Fretheim (2006 : 182), citant Gundel (1999), affirment que « *all sentences have an information focus, as an essential part of the function of sentences ininformation processing; but not all sentences/utterances have a contrastive focus, the latter being determined primarily by a speaker/writer's intention to affect the addressee's attention state at a given point in the discourse* ».

Enfin, mentionnons Gundel (1999 : 293) qui propose trois types de focus.

- **focus psychologique** ou **d'attention** : « *an entity is in (psyhocological) focus if the attention of both speech participants can be assumed to be focused on it because of its salience at a given point in the discourse* » (Gundel 1999 : 294).

- **focus linguistique :** constitué de deux sous-types

  - **sémantique** ou **d'information**, « *that part of a sentence that is prosodically (and sometimes also syntactically) prominent,* […] *that answers the relevant wh-question (implicit or explicit) in the particular context in which the sentence is used* » (Gundel 1999 : 295), « *information that is being predicated in relation to the topic* » (Gundel 1999 : 298). Il s'agit donc de la réponse à une question partielle ou bien d'une information nouvelle affirmée du topique. Il est

[1] Charolles (2002 : 92) parle de relation de nature *partonomique* lorsque le référent comporte « des parties qui en constituent des composants bien connus ».

le corrélat linguistique de la notion pragmatique de 'commentaire' (Gundel 1999 : 299, 302).

- **focus contrastif**, contrasté explicitement ou implicitement avec quelque chose d'autre, « *imposed salience on a given constituent (often, but not necessarily, a topic) that the speaker wants to call the addresse's attention to in order to contraste it with other potential members of a relevant 'contrast set'* » (Gundel 1999 : 298). Il s'agit d'une saillance mise sur un constituant pour attirer l'attention, introduire un nouveau topique, et effectuer un contraste.

Gundel (1999) conçoit les focus sémantique et contrastif comme porteurs d'un *prominent pitch accent* (Gundel 1999 : 298) permettant d'attirer l'attention sur le constituant concerné. Le focus d'information est typiquement voire universellement mis en relief grâce à la prosodie (Gundel et Fretheim 2006 : 181). L'association entre la proéminence prosodique et le focus a en effet été étudiée et démontrée pour une grande variété de langues typologiquement et génétiquement différentes (Gundel et Fretheim 2006 : 184). Gundel (1999 : 303) précise en note que le focus sémantique peut également être marqué syntactiquement.

Les deux focus linguistiques impliquent tous deux un effet de contraste : il s'agit de la fonction primaire du focus contrastif, mais seulement d'un effet secondaire du focus sémantique. Le focus contrastif se distingue donc du focus sémantique par sa fonction primaire. Celle du focus contrastif est de marquer le contraste ou l'emphase (Gundel 1999 : 299-300). Celle du focus sémantique est de marquer l'information nouvelle prédiquée par le topique. Les constituants qui reçoivent un focus contrastif « *are emphasized because they reprensent the new information being predicted of the topic. The fact that this new information is (implicitely or explicitely) in contrast with other things that may have been predicated of the topic is a secondary effect* » (Gundel 1999 : 298). On voit qu'elle tisse un lien entre structure et statut de l'information. Seul le focus sémantique est capable de modifier le sens de la phrase et d'avoir un effet sur ses conditions de vérité (Gundel 1999 : 302). « *This is because information focus is a relational notion that determines the main predication in the sentence, that predication being assessed relative to the topic. Purely contrastive focus has no truth-conditional effects* » (Gundel et Fretheim 2006 : 190).

Les trois types de focus sont indépendants du point de vue logique mais ne s'excluent pas. Ils peuvent être présents dans une même phrase, voire sur un même constituant. Un constituant dont le référent ne fait pas partie du focus psychologique peut faire partie du focus sémantique, mais pas nécessairement (Gundel 1999 : 299). Le focus sémantique est complémentaire du focus psychologique et ne suffit pas, en soi, à mettre une entité dans le focus psychologique (Gundel 1999 : 300), à moins qu'il ne soit marqué à la fois syntactiquement et prosodiquement (Gundel 1999 : 303). Au contraire, comme la fonction principale du focus contrastif est d'attirer l'attention sur quelque chose, le focus contrastif déclenche toujours un focus psychologique (Gundel 1999 : 301).

3.5. Dans le syntagme nominal

Toute focalisation a lieu à l'intérieur d'un segment bien précis de l'énoncé, nommé domaine de focalisation chez (Nølke 2001 : 95). Le syntagme nominal peut-il être un tel domaine ? Un certain

nombre d'auteurs ont supposé que oui. Nous mentionnons ici deux courants théoriques différents, mais qui s'accordent quant au fait que le focus dépend de l'ordre des mots, donc de la syntaxe. Ils portent sur la place de l'adjectif dans le syntagme nominal dans les langues romanes. De nombreux travaux existent sur la position de l'adjectif en français, mais ceux de Nølke permettent de combiner des facteurs syntaxiques, sémantiques et formels pour décrire et prédire les faits.

1) Tout syntagme nominal constitue un domaine de focalisation mineure (Nølke 2001 : 180). L'auteur décrit ainsi l'organisation du syntagme nominal français :

> « Antéposition : (Adj (Sub))
> Postposition : (Sub (Adj))
>
> Découlent directement de ces structurations les deux règles suivantes :
> - Un adjectif antéposé ne constitue jamais à lui seul le foyer simple : ou bien il est focalisé avec son substantif, ou bien il se trouve en dehors du foyer.
> - Un adjectif postposé est toujours focalisé : ou bien il forme le foyer avec son substantif, ou bien il est seul à être focalisé. » (Nølke 2001 : 180)

Antéposé, l'adjectif est mis au premier plan *avec* le nom, ce qui produit un effet de mariage sémantique (Nølke 2001 : 181). Lebas-Fraczak (2015 : 40) explique qu'« un adjectif antéposé peut ou bien se trouver "en dehors du foyer", c'est-à-dire ne pas être focalisé, ou bien être focalisé avec son substantif, tandis que l'adjectif postposé est toujours focalisé, soit seul soit avec son substantif ». Elle avance cependant que « dans le cas d'un adjectif postposé, la postposition ne peut pas mettre le nom complètement "en dehors du foyer" » (Lebas-Fraczak 2015 : 40).

Une telle organisation du syntagme explique pourquoi le sens de l'adjectif semble inhérent ou peu spécifique lorsqu'il est antéposé, et pourquoi il peut avoir une valeur spécifique lorsqu'il est postposé. C'est que le sens spécifique permet d'établir des distinctions, donc d'identifier le référent du syntagme tout entier. « Plus un adjectif a pour fonction de dégénéraliser le sens du substantif plus l'adjectif a l'occasion d'être postposé. Moins un adjectif dégénéralise plus un adjectif a l'occasion d'être antéposé » (Nølke 2001 : 172). Autrement dit, l'emplacement de l'adjectif dépend de sa sémantique. Plus un adjectif est spécifique, plus il comporte de sèmes spécifiques, moins son extension est grande : il a donc plus de chance d'être postposé car il favorise la focalisation (Nølke 2001 : 181, 190). Moins un adjectif est spécifique, moins il comporte de sèmes spécifiques, plus son extension est grande et son sens général : il a plus de chance d'être antéposé car il défavorise la focalisation. Or « la focalisation s'appuie sur les sèmes spécifiques » (Nølke 2001 : 182), et l'antéposition est l'emplacement le moins focalisé. L'emplacement de l'adjectif dépend donc de sa sémantique : son degré de spécificité et son usage le rendent plus ou moins facilement postposable ou antéposable.

- les adjectifs dits « élémentaires » (*bon, grand, gros*) ont tendance à être antéposés (Nølke 2001 : 186, 190).
- Les adjectifs qui « s'approche[nt] des déterminants » (*pure bêtise*) et les numéraux ordinaux (*premier, deuxième*) sont antéposés (Nølke 2001 : 186).

- Les adjectifs exprimant la quantification (*nombreux, vaste*, mais aussi *premier, deuxième*, *dernier*) sont antéposés (Nølke 2001 : 186-187).
- un adjectif exprimant une propriété inhérente au nom (concernant sa taille, sa quantité, ses relations à l'univers du discours) est antéposé. Il a perdu son éventuelle fonction classificatrice car les sèmes spécifiques, absents de l'antéposition, servent à classifier. En effet, une propriété inhérente au nom n'implique pas de choix du référent à l'intérieur d'un paradigme.
- un adjectif employé de manière métaphorique ou comme épithète de nature est antéposé car ses sèmes spécifiques sont atténués (Nølke 2001 : 185-186).
- un adjectif de faible extension, lorsqu'il est antéposé, perd ses sèmes spécifiques et gagne en extension. Cela permet de lui attribuer diverses valeurs, connotations et effets de style. S'il faut employer un adjectif peu spécifique d'une manière spécifique, il sera postposé :

> « Par là, les sèmes (les traits sémantiques) spécifiques qui se trouvent dans le foyer seront soulignés, alors que les sèmes spécifiques se trouvant hors du foyer seront atténués. Autrement dit, le lexème focalisé garde son sens plein, alors que le lexème non focalisé perd de sa valeur sémantique. » (Nølke 2001 : 180)
>
> « Inversement, dans l'antéposition, nous en avons vu de nombreux exemples, un adjectif normalement postposé perd de ses sèmes spécialisés, ce qui peut avoir une large gamme d'effets divers (valeur subjective, connotative, banale, anaphorique, métaphorique, d'épithète de nature, adverbiale, …). » (Nølke 2001 : 191)

- un adjectif précis de couleur (*doré*, *rosâtre*), porteur de nombreux sèmes, partant plus spécifique, peut difficilement être antéposé, car il favorise la focalisation – contrairement à un adjectif de couleur très global, porteur de moins de sèmes (*jaune*, *rouge*).
- un adjectif relationnel n'accepte pas l'antéposition (Nølke 2001 : 190-191) lorsqu'il est employé au sens propre.

Nølke (2001 : 199, 209) en tire une topologie du syntagme nominal en français (Figure 19) :

Déterminant	Pré-zone	Substantif	Post-zone

Figure 19 : Topologie du syntagme nominal (Nølke 2001 : 199, 209)

La pré-zone forme quasiment toujours un seul intonème avec le déterminant et le substantif, et ne permet pas la présence d'autres modifieurs que les adjectifs épithètes à fonction *modificatrice* (Nølke 2001 : 199, 210-211), c'est-à-dire graduables. Ils ont un sens « atténué et plus ou moins général, et il est souvent question d'une fusion sémantique et prosodique (la liaison est habituelle) avec le substantif » (Nølke 2001 : 200). La post-zone contient des épithètes à fonction modificatrice ou bien *relationnelle*, c'est-à-dire non-graduables (adjectifs relationnels, qu'il nomme pseudo-adjectifs), mais aussi des appositions, des syntagmes prépositionnels et des subordonnées relatives (Nølke 2001 : 210). L'impossibilité d'antéposer un adjectif relationnel « s'explique par le fait qu'[il] ne renferme

aucun sème générique qui pourrait être mis en valeur dans l'antéposition : en effet, son seul sème est le seul relationnel »[1].

2) Cinque (2010, 2014), s'inspirant notamment de Sproat et Shih (1990), pose lui aussi deux types distincts de position adjectivale. Ils reflètent la distinction majeure suivante :

- **modification directe** : lecture ponctuelle, non-restrictive, modale, non-intersective, absolue de l'adjectif et du superlatif, impliquant la spécificité, évaluative, dépendant du syntagme nominal,
- **modification indirecte** : lecture générale, restrictive, comme une proposition subordonnée relative implicite, intersective, relative à une classe de comparaison, comparative du superlatif, n'impliquant pas de spécificité, épistémique, anaphorique du discours.

Les adjectifs qui sont des modifieurs indirects sont, à l'origine, les prédicats de propositions relatives réduites (*reduced relative clause*). Les adjectifs qui sont des modifieurs directs sont, à l'origine, non-prédicatifs et « *are introduced as specifiers of functional projections within the extended projection of the noun* » (Alexiadou 2014 : 11). Il existe donc deux approches différentes de la modification adjectivale. En français, on distingue modification à gauche (« modification directe » de l'adjectif anténominal) et modification à droite (« modification indirecte » de l'adjectif postposé), semblable à une proposition relative.

> « L'adjectif épithète antéposé est donc adjacent au N, tandis que l'épithète postposé est plus mobile. En termes structuraux, l'adjectif antéposé est toujours adjoint au N, tandis que l'épithète postposé est soit au même niveau que les compléments de N (avec lesquels il permute), soit adjoint au GN. Si l'on reprend la distinction (à vocation universelle) proposée par Sproat et Shih (1988) entre modification directe et indirecte, on peut dire que l'épithète antéposé est toujours un modifieur direct, tandis que l'épithète postposé a les deux possibilités. » (Abeillé et Godard 1999 : 16)

Cinque (2010, 2014) choisit une perspective plus large et compare les langues germaniques aux langues romanes (dont le français). Il prend pour exemple l'anglais (5) et l'italien (6).

(5) Anglais (Cinque 2010 : 92)

		Indirect mod.	Direct mod.	Direct mod.	
a.	*the*		*American*	*musical*	*comedy of the 1950s*
b.	*the only*		*possible*	*Roman*	*invasion of Thrace*
c.	*She interviewed every*	*possible*	*potential*		*candidate*

[11] Cela n'empêche pas de pouvoir, dans certains contextes, antéposer un adjectif relationnel : *le très catholique Mr Dupont*. L'adjectif change alors de sens et perd sa propriété relationnelle initiale. Il prend un nouveau sème, à valeur graduable, émotionnelle, éventuellement connotative (Nølke 2001 : 211).

(6) Italien (Cinque 2010 : 92)

		Direct mod.	Direct mod.	Indirect mod.	
a.	*la commedia*	*musicale*	*americana*		*degli Anni '50*
b.	*la sola invasione*	*romana*	*possibile*		*della Tracia*
c.	*Intervistò ogni candidato*	*potenziale*		*possibile*	*alla presidenza*

Cinque (2014 : 3) analyse les faits comme suit :

> « *We will observe the general pattern in (3). English adjectives are ambiguous in pre-nominal position, while Italian adjectives are ambiguous in post-nominal position. English adjectives in post-nominal position, when possible, are instead unambiguous and so are the pre-nominal adjectives in Italian (which furthermore take the opposite values of the English post-nominal adjectives).*
>
> (3) a. English: AP N (AP)
> (**ambiguous**) (**unambiguous**)
> b. Italian: AP N AP
> (**unambiguous**) (**ambiguous**) » (Cinque 2014 : 3)

Il constate que ce qui se produit en anglais se produit également en allemand : dans les langues germaniques, le modifieur indirect précède le modifieur direct en position anténominale. Ce qui se produit en italien se produit également en français et en espagnol : dans les langues romanes, le modifieur indirect suit le modifieur direct en position postnominale. Cela donne une structure en miroir, ainsi que le montrent les Figure 20 et Figure 21.

Langues germaniques :	DEF A N =	modification directe ou indirecte
	DEF N A =	modification directe
Langues romanes :	DEF A N =	modification directe
	DEF N A =	modification directe ou indirecte

Figure 20 : Schéma résumant les analyses de Cinque (2010)

English (Germanic)
AP from reduced RC > 'direct modification' AP > N > AP from reduced RC
Italian (Romance)
'direct modification' AP > N > 'direct modification' AP > AP from reduced RC

Figure 21 : Résumé de l'ordre des adjectifs dans un syntagme nominal en anglais et en italien (Cinque 2010 : 22)

Chapitre III : État de l'art

Ce chapitre présente la structure qui fait l'objet de notre étude. Il introduit tout d'abord le syntagme nominal en romani : les parties du discours principales qui entrent dans sa composition (nom, déterminants, adjectif) et l'ordre dans lequel on les emploie. Il fait ensuite état des auteurs qui ont relevé et tenté d'expliquer le phénomène de répétition de l'article défini en romani. Enfin, il présente les études consacrées à une structure qui semble identique, le « syntagme polydéfini », abondamment décrit et théorisé pour une autre langue, le grec moderne.

1. Le syntagme nominal en romani

On appelle syntagme nominal un syntagme dont la tête est un nom.

1.1. L'ordre des constituants

L'ordre canonique des constituants nominaux en romani est, selon Matras (2002 : 166) – qui choisit d'y inclure la préposition introduisant éventuellement le syntagme :

> « *Table 7.1 Linear layout of the noun phrase: principal slots*
>
> *[preposition] + [determiner] + [quantifier] + [adjective] + noun + [options]* »

On voit qu'à gauche de la tête se trouve la quasi-totalité des éléments constitutifs du syntagme nominal : déterminant, quantificateur, adjectifs, tous porteurs, sauf exception, de marques de genre, de nombre et de cas. À droite de la tête se trouvent les « options » : syntagme génitival, syntagme prépositionnel, adverbe, proposition relative, apposition. L'exemple (7) présente plusieurs constituants dans l'ordre canonique : déterminant (possessif) + adjectif + nom + option (ici une proposition relative).

(7) *Romani (F60A, Korçë, 21 juillet 2014)*

Sora	*isi*	*m-o*	*bar-o*	*phral/*	*kaj*	*čard-ena*	*les*	*Yli.*
après	être.3.PRES	POSS.1SG-M.SG	grand-M.SG	frère(M.SG)	REL	appeler-3PL	3SG.OBL	Yli

« Ensuite il y a mon grand frère/ qui s'appelle Yli. »

On constate que, d'une part, il n'y a de place que pour un seul déterminant, qui occupe la première place du syntagme nominal après une éventuelle préposition, et que, d'autre part, l'adjectif se place avant le nom. Plus précisément, on trouve dans le syntagme nominal :

1) un **déterminant** : article défini (fusionnant éventuellement avec la préposition), article indéfini, zéro, démonstratif, possessif, quantifieur, numéral, syntagme génitival,
2) un **quantifieur** ou un **numéral** – qui revêtent la fonction d'un déterminant ou d'un adjectif.
3) un **adjectif :** par exemple *lolo*, rouge, éventuellement dénominal, hérité ou emprunté. Le syntagme génitival peut également remplir ce même rôle.
4) un **nom.**
5) un **élément optionnel** : adjectif (section 1.4.3, p. 138), syntagme génitival, proposition relative…

Masica (1991 : 373) propose le schéma suivant pour les langues indo-aryennes modernes :

{Gen P}

+ Det + (Quant) + (Adj P) + N

{Loc P}

Ce schéma correspond à celui de Matras (2002 : 166) pour le romani, à ceci près que les « options » ne sont pas mentionnées, et que le syntagme génitival est antéposé au déterminant. Le déterminant n'occupe donc pas la première position absolue du syntagme dans les langues indo-aryennes autres que le romani : le syntagme génitival (*GenP*) ou le syntagme locatif (*LocP*) peuvent le précéder.

1.2. Le nom

1.2.1. Catégories

Le nom présente les principales catégories des langues indo-européennes. Les catégories grammaticales et sémantiques, propres au lexème nominal, sont les suivantes en romani :

- **Genre** lexical : masculin et féminin.
- **Animéité** : animé et inanimé. On peut parler d'une hiérarchie de l'animéité en romani (Boretzky 1994a : 101) : la catégorie animée comprend les êtres humains et les animaux considérés comme supérieurs, la catégorie inanimée comprend les animaux considérés comme inférieurs, les objets, les concepts. La répartition des animaux entre l'une et l'autre classe se fait selon leur plus ou moins grande proximité avec l'homme, et cette répartition diffère selon les variétés. Pour une étude approfondie sur la question de la distribution des cas en fonction des référents et des catégories lexicales, on pourra consulter Matras (1997 : 72-78), la hiérarchie détaillée proposée (Matras 1997 : 77) et la synthèse sur la hiérarchie de *retrievability*, récupérabilité (Matras 1997 : 78). Ces distinctions sont grammaticalisées dans

la morphologie de l'objet des verbes transitifs, marqué aux cas oblique pour un « animé » et direct pour un « inanimé », ainsi que dans l'usage du paradigme anaphorique, les « animés » étant repris par des pronoms personnels et les « inanimés » par des démonstratifs (Matras 2002 : 72-73).

- **Statut thématique** : thématique et athématique. Cette catégorie concerne la morphologie des noms et des adjectifs (Matras 2002 : 73). Sont considérés thématiques les éléments hérités d'origine indo-aryenne et dravidienne, mais aussi iranienne, arménienne et grecque. Il s'agit du vocabulaire commun aux ancêtres des Roms avant qu'ils ne se dispersent en Europe, donc du romani ancien. Sont considérés athématiques les éléments des langues de contact postérieur, souvent empruntés avec les phonèmes étrangers et l'accent tonique non-final. Ces éléments peuvent être intégrés au système linguistique du romani grâce à des morphèmes, pour la plupart d'origine grecque, qui les « thématisent » (*-i* final pour les noms…) ou non (*-ime* final pour les participes ou les adjectifs…). Selon Matras (2002 : 74), la scission est si forte que l'on peut parler d'une « grammaire athématique » :

> « *The fact that we are dealing with a distinction that manifests itself in more than just one grammatical domain justifies the notion of an athematic 'grammar'* […]. *Moreover, there is evidence that the Greek morphology that was adopted into Romani for the purpose of loanword integration in fact constituted a pattern of morphological adaptation that was applied to words of foreign origin already in Greek. Romani may therefore be said to have borrowed a Greek borrowing pattern (Bakker 1997b). Although all Romani dialects show a thematicity split, many have replaced some of their nominal athematic morphology of Greek origin through later loan morphology, especially nominative plural endings*. » (Matras 2002 : 74)

Les catégories syntaxiques, qui témoignent de la relation entre le lexème nominal, le syntagme nominal, et les autres syntagmes et constituants de la phrase, sont les suivantes :

- **Nombre** : singulier et pluriel, qui correspondent à des valeurs sémantiques d'unicité et de paucalité/pluralité (sauf dans le cas des noms indénombrables, morphologiquement au singulier). Pour les adjectifs, pronoms et démonstratifs, les genres sont morphologiquement distincts au singulier mais indistincts au pluriel.
- **Cas** : deux cas fusionnels (direct, oblique), cinq cas agglutinants (datif, locatif, ablatif, instrumental, génitif) ainsi que le vocatif. Nous employons le terme de *cas* dans un sens strictement morphologique (*oblique* étant un état dans lequel se trouvent les noms, qui prennent une *désinence* particulière), et non dans un sens fonctionnel (on parlera d'*objet*, de *sujet*…).
- **Définitude** : défini et indéfini (voir ci-dessous). La définitude joue un rôle dans le marquage différentiel de l'objet dans certaines variétés[1] où les êtres animés définis sont marqués au cas oblique tandis que les être animés indéfinis sont au cas direct.

Le Tableau 10 synthétise les catégories du nom en romani.

[1] cf. Boretzky (1993 : 22) pour le bugurdži du Kosovo, Boretzky (1994a : 98, 101-102) pour le kalderaš du Banat, Matras (2002 : 86) pour la variété des Polska Roma et Cech (2006 : 22) pour le dolenjski de Slovénie.

Tableau 10 : Synthèse des catégories nominales

Genre	Animéité	Statut thématique	Nombre	Cas	Définitude
masculin	animé	thématique	singulier	direct	défini
féminin	inanimé	athématique	pluriel	oblique	non-défini
				vocatif	
				datif	
				locatif	
				ablatif	
				instrumental	
				génitif	

1.2.2. Flexion

Le système de flexion des noms et pronoms du romani comporte trois strates (*layers*), comme les autres langues indo-aryennes modernes (Masica 1991 : 230-248) :

1) **une première strate synthétique fusionnelle** constituée des cas polyvalents direct et oblique[1]. Ils sont hérités des cas fusionnels du moyen-indien et portent donc l'accent final de mot. Le cas direct marque le sujet des phrases, le complément des prépositions, l'objet direct non « animé ». Le cas oblique est utilisé pour l'objet direct « animé », l'amarrage des suffixes agglutinants de la deuxième strate, l'expression de la possession. Selon Matras (2002 : 86), les cas direct et oblique se répartissent en fonction du degré de topicalité du référent. Comme nous l'avons mentionné plus haut, le marquage différentiel de l'objet (cas direct ou oblique) est fonction de l'animéité du référent, selon une échelle complexe et variable, mais aussi de sa définitude et de sa spécificité. Il pourrait s'agir d'un héritage : la définitude ou la familiarité est également un facteur de marquage différentiel de l'objet en domari, langue indienne cousine parlée au Proche-Orient (Matras 1999b, Matras 2012 : 143, Herin 2012 et Herin 2014).
2) **une deuxième strate synthétique agglutinante** constituée de cas monovalents : datif, locatif, ablatif et instrumental. Ils sont invariables en genre et en nombre. À ces cas s'ajoute le génitif, dont une partie est invariable (morphème court *-k-* ou long *-ker-*[2]) et l'autre variable en genre et en nombre (*-i, -o, -e*) en fonction du nom antécédent du syntagme génitival. Les suffixes de la strate agglutinante s'accolent au nom ou pronom marqué au cas oblique, et ne peuvent porter l'accent tonique. Notons que les déterminants et modifieurs qui portent sur le nom fléchi à la strate 2 ne se trouvent pas eux-mêmes fléchis suivant la strate 2 mais suivant la strate 1, au cas oblique.
3) **une troisième strate analytique** qui exprime la fonction sémantique du nom au moyen de prépositions. Celles-ci sont héritées du moyen-indien (*k*, à, *an(d)*, pour) ou empruntées tardivement (*ku* et *me*, avec, empruntés respectivement au roumain et au grec/à l'albanais). Elles sont d'origine déadverbiale (*tel*, sous, de *tele*, en-dessous) ou dénominale (*maškar*, entre, parmi, de *maškar*, centre, milieu). C'est la strate la plus récente, mais aussi la plus productive

[1] Nous renonçons aux termes *nominatif* et *accusatif* afin de ne pas mettre sur le même plan les cas flexionnels de la strate 1 et ceux, agglutinants, de la strate 2 (génitif, locatif…).

[2] Et leurs allomorphes *-g-* et *-ger-*.

et la plus sujette au renouvellement. Selon Boretzky (1993 : 38), les prépositions peuvent remplacer progressivement les cas synthétiques de la strate 2. Matras (1997 : 78-88) analyse en détail le statut et la solidité des différents cas des strates 1 et 2, proposant la hiérarchie de stabilité suivante (Matras 1997 : 88) : (NOM) > GEN > INSTR > DAT > ACC > ABL > LOC. Il montre que ce phénomène dépend notamment des variétés et propose une hiérarchie entre cinq variétés jugées représentatives des cinq grands groupes dialectaux. L'érosion des cas synthétiques est par exemple peu avancée en Lituanie (Tenser 2005 : 40). Elle dépend en outre des ressources de vocabulaire disponibles, des cas et des substantifs (Matras 2002 : 92-93).

1.3. Le déterminant

Le romani présente deux articles (défini et indéfini) ainsi que des pronoms (démonstratif, possessif) et des quantificateurs (quantifieur, numéral) qui peuvent faire fonction de déterminant, de même que le syntagme génitival. À l'exception des quantificateurs, ils encodent la définitude, au contraire de l'article indéfini et de l'article zéro (Halwachs et al. 2002 : 48).

1.3.1. Article défini

1.3.1.1. Origine

La définitude est marquée en romani par un article défini distinct du démonstratif (Dryer 2013). Le romani est la seule langue indienne moderne à présenter un article défini totalement grammaticalisé. On a longtemps cru que le romani l'avait emprunté au grec, en raison de son homophonie avec les articles grecs masculin [o] et féminin [i] au nominatif singulier (Paspati 1870 : 39). Le reste du paradigme montre cependant qu'il n'en est rien. Comme dans de nombreuses langues du monde, il provient de l'ancien démonstratif, de même que le pronom 3SG et certaines formes de la copule. Boretzky (2000b : 54-59) propose une reconstruction des formes historiques, et Matras (2002 : 106-112) propose un cycle de mutation et de renouvellement de ces éléments, à travers trois stades précis d'évolution en proto-romani et un stade pour le romani ancien (*Early Romani*).

Si le romani n'a pas emprunté au grec les formes de l'article défini, on peut dire qu'il lui a emprunté l'idée, le principe d'article défini et la plupart de ses emplois (cf. section 1.3.1.4, p. 111). En effet, toutes les variétés du romani connaissent ou ont connu un article défini, et des formes sont aujourd'hui encore communes à toutes les variétés – ce qui d'ailleurs confirme qu'il est né à une date très ancienne, avant la dispersion des locuteurs à travers l'Europe, à l'époque du romani commun où l'influence hellénique était très forte. Il est possible que le grec ait également influencé la réduction phonologique qui s'est opérée, les locuteurs rapprochant l'article des deux langues vers des formes homophones [o] et [i] au cas sujet singulier, masculin et féminin (Matras 2002 : 112). Boretzky (2000b : 56) avance toutefois que la forme féminine *e* serait plus ancienne que la forme *i*, parce que l'origine de la première est inexplicable, tandis que la seconde s'explique par la finale -*i* des noms

(*gurumni*, vache), adjectifs (*parni*, blanche) et de certains pronoms (*kasavi*, telle) féminins. La réduction phonologique a dû se produire très tôt, dès le romani ancien, car les deux formes *o* et *i/e* sont les deux seules communes à toutes les variétés : les formes direct pluriel et obliques, auraient fait l'objet d'une réduction moindre, plus tardive, et distincte selon les variétés. Boretzky (2000b : 58) donne d'autres arguments pour une origine grecque de l'article défini : le romani aurait emprunté à l'article défini grec sa position et sa structure phonologique.

> « […] *there were no other languages in the Near East with ART or similar article behavior, whereas the numerous similarities between Greek and Romani ART leap to the eye: first position in the NP, mostly vocalic in the nominative, mostly CV in the oblique cases, no stressed forms, no reinforcement in the bulk of the dialects, no isolated (demonstrative) use, use with personal names even in dialects remote from the Greek area. One might argue that* ART *may have come up in a totally internal process, but though there is no logical argument excluding it, little speaks in favor of it, first of all because true ART languages are rare among the languages of the world, and further because there seem to be few isolated ART languages around the world, for which isolated development of this element must be presupposed. The definite article is an areal phenomenon.* » (Boretzky 2000b : 58)

Boretzky (2000b : 55, 58) rejoint Lyons (1999) et Dryer (2013), selon lesquels l'expression de la définitude via un article défini constitue un phénomène aréal. Cette position est d'ailleurs confirmée par le fait que le romani ait perdu l'article défini dans les variétés en contact avec des langues sans article défini (Matras 2000a : 183) : celles d'Iran (zargari, romano), de Russie du nord, de Pologne du sud (lovari), de Finlande (Boretzky 2000b : 59), de Pologne (Polska Roma, xaladitka), de Croatie (hrvati) (Matras 2002 : 96), de Slovénie (dolenjski) (Cech 2006 : 20), de Lituanie (Tenser 2005 : 24)[1]. La perte est partielle dans les variétés du nord-est selon Boretzky (2000b : 34) ; on y constate parfois un usage résiduel avec les prépositions amalgamées et avec les noms propres. Comme le fait remarquer Matras (2002 : 96), c'est une preuve supplémentaire du caractère aréal de l'article défini : le romani a acquis un article défini par contact dans une région où toutes les langues présentent un article défini, et ses variétés l'ont ensuite perdu dans les régions où les langues n'en présentent pas. La perte de l'article défini (Matras 2000a : 183) est l'un des symptômes de la « dé-balkanisation » des variétés du romani hors des Balkans (Matras 2000a : 190). Ce développement est un cas de *replacive convergence*, qu'il définit comme « *a new layer of convergent structures that replaces the older layer of Balkan convergent features* » (Matras 2000a : 182). Certaines variétés échappent à ce phénomène : Adamou (2016 : 158-159) constate qu'en ce qui concerne l'article défini, le romani de Thrace n'a pas convergé avec sa langue dominante de contact, le turc. En effet, le turc n'a pas d'article défini et le romani est parvenu à conserver le sien ; sur ce plan, il présente donc de plus de similarités avec le grec, sa deuxième langue de contact.

[1] Tenser (2005 : 24) note néanmoins que l'article peut être employé optionnellement aux cas direct et oblique.

1.3.1.2. Flexion

Le romani appartient au type 1 de la typologie de Dryer (2013) : l'article défini y est distinct du démonstratif. Il se fléchit en genre, nombre et cas – avec une grande variation dialectale, notamment pour les formes obliques (Matras 2002 : 97) :

Tableau 11 : Morphologie de l'article défini dans les groupes de dialectes du romani (Matras 2002 : 97)

	NOM			OBL		
	SG.M	SG.F	PL	SG.M	SG.F	PL
Northeastern	*o*	*i/e*	*o*	*e*	*e*	*e*
Northwestern	*o*	*i/e*	*i/o*	*i/e*	*i/e*	*i/e*
Northern Central	*o*	*e*	*o*	*(l)e*	*(l)a*	*(l)e*
Southern Central	*o*	*i*	*o*	*(o)le*	*(o)la*	*(o)le*
Vlax	*o*	*i/e*	*(l)e/əl/ol*	*(l)e*	*(l)a/e*	*(l)e/o(l)*
Bugurdži group	*o/u*	*i*	*o/u*	*e*	*e/i*	*e*
Southern Balkan	*o*	*e/i*	*o*	*e*	*e*	*e*

Les formes que j'ai pu entendre en Albanie sont présentées dans le Tableau 12. L'article défini présente en arli et mečkar d'Albanie les mêmes formes qu'en arli de Macédoine, tandis qu'il présente en čergar I les mêmes formes qu'en gurbet de Macédoine (Friedman 1993 : 19).

Tableau 12 : Morphologie de l'article défini des trois principales variétés d'Albanie

		Singulier		Pluriel
		Masculin	Féminin	[Masc=Fém]
Cas direct	mečkar & arli	*o*	*i*	*o*
	čergar I	*o*	*e*	*e*
Cas oblique	mečkar, arli, čergar I	*e*	*e*	*e*

1.3.1.3. Fusion avec certaines prépositions

Certaines prépositions héritées incorporent l'article défini (Figure 22). Cette fusion rappelle ce qui se passe dans les langues romanes (français *au* < *à* + *le*, italien *al* < *a* + *il*), en allemand (*im* < *in* + *dem*) ou en grec (στο < σε + το) avec les amalgames préposition + article défini.

ka/ke	+	*O*	=	***ko***	
à		ART.DEF.M.SG		à.ART.DEF.M.SG	= « au »
pe	+	*i*	=	***pi***	
sur		ART.DEF.F.SG		sur.ART.DEF.F.SG	= « sur la »
ande	+	*i*	=	***andi***	
dans		ART.DEF.F.SG		dans.ART.DEF.M.SG	= « dans la »
paše	+	*o*	=	***pašo***	
près		ART.DEF.M.SG		près.ART.DEF.M.SG	= « près du »

Figure 22 : Fusion de l'article défini avec une préposition

Comme le note Matras (2002 : 165), il y a homophonie, donc ambiguïté possible, lorsque l'article défini a lui-même la forme *e* (FEM.SG ou PL selon les variétés). Les syntagmes (10) et (11) sont homophones : (10) est indéfini mais (11) est défini.

(8)	*ande*	*jekh them*	
	dans	IND pays	
	« dans un pays »		= préposition + article indéfini
(9)	*ando*	*them*	
	dans.ART.DEF.M.SG	pays	
	« dans le pays »		= préposition incorporant l'article défini
(10)	*ande*	*them-a*	
	dans	pays-M.PL	
	« dans des pays »		= préposition + article zéro
(11)	*ande*	*them-a*	
	dans.ART.DEF.M.PL	pays-M.PL	
	« dans les pays »		= préposition incorporant l'article défini

Adamou (2016 : 158) note l'existence de la fusion de l'article défini avec plusieurs prépositions en romani de Thrace : « andi jak *'in the fire'* », et précise que ce phénomène existe aussi en grec « sti fotia *'in the fire'* ». Pour Boretzky (1994a : 93), il n'y a pas seulement fusion entre les deux éléments, mais un processus de *morphologisation* amenant les prépositions à devenir flexionnelles. Ce processus existerait dans toutes les variétés, quelle que soit la forme qu'aient prise les prépositions héritées (*ando* ou *dro*, dans.ART.DEF.M.SG) : il serait favorisé par leur grande fréquence d'emploi et par l'impossibilité de faire fusionner l'article avec le nom qui suit. Le processus est allé si loin que certaines prépositions ne peuvent plus, selon Boretzky (2000b : 44), s'employer sans l'article fusionné. Dans la pratique, la préposition est très rarement suivie de l'article indéfini (Boretzky 2000b : 44) :

> « *So far we have nearly no knowledge of the details, but we can state that there is a preference for the definite article, especially with prepositions ending in a vowel usually deleted before the (vocalic) article. In Kalderaš, one might expect to hear* ande gav *"in a village" or* ke gav *"to a village" besides the frequent* and-o gav *and* k-o gav*, but I have a strong feeling that the constructions without ART are highly marked and do not appear in normal discourse. What is more, in some dialects there is a strong tendency to keep ART even when the PP contains an additional determiner, e.g. Bugurdži* k-o mo phral *rather than* ke mo phral *"to my brother", and definitely* k-o than leskoro *"in his stead", or Arli* an-o lesoro sanduko *"in his coffin", and* tar-i odoja crkva *"from that church" (Boretzky 1996:7). In such cases the ART tends to turn into something like a gender marker for prepositions*. » (Boretzky 2000b : 44)

Le romani suit ici un processus typologiquement commun, qui fait ultimement de l'article défini un simple marqueur de nominalité ou de genre (cf. Greenberg 1978 : 69). Si l'article a pris la fonction de marquage de genre (et de nombre ?) de la préposition, il peut alors avoir perdu le trait de définitude. Celui-ci est en effet déjà porté, dans ses exemples, par les définis complexes possessifs et

démonstratifs. Dans les variétés qui ont perdu l'usage de l'article défini (cf. section 1.3.1.1, p. 107), certaines prépositions semblent s'accorder en genre et en nombre avec le nom qui suit : il s'agit du seul résidu attestant de l'existence d'un ancien article défini (Ackerley et Winstedt 1925 : 76, Matras 1999a : 9-10, Tenser 2005 : 24). Ce fait peut être interprété comme une preuve supplémentaire du haut degré de morphologisation de la préposition.

1.3.1.4. Distribution

L'article occupe la première position dans le syntagme nominal, précédant toujours le nom qu'il détermine. Matras (2002 : 165-166) précise que les déterminants assignant la définitude sont incompatibles entre eux : démonstratifs, interrogatifs, possessifs, articles définis. Boretzky (2000b : 41) mentionne en revanche des cas de « double détermination » où le syntagme présente deux déterminants, par exemple article défini et démonstratif.

1.3.1.5. Fonctions

Les contextes d'emploi sont globalement les mêmes que dans les autres langues indo-européennes d'Europe et notamment des Balkans. Boretzky (2000b : 45-53) dresse une typologie des usages de l'article, dont la fonction principale de l'article défini est d'établir l'unicité de la référence. Les fonctions présentées par Boretzky (2000b : 44-54) sont résumées dans le Tableau 13.

> « *We will follow Himmelmann (1997) who, elaborating on Hawkins (1978), Loebner (1985) and others, distinguishes pragmatic from semantic uses, the anaphoric ones being but one among many others. Frequency counts for Western European languages point to the fact that ART occurs more often with non-anaphoric functions, but this cannot be confirmed for Romani in general. However, the reason for this divergence may be caused by the kind of texts analysed: whereas for other languages it may be everyday discourse that was analysed the texts available in Romani are mainly compact stories told by one person: fairy tales, descriptions of remarkable events, anecdotes and the like, where immediate reference to actual situations is possible in direct speech but hardly elsewhere. My impression is that anaphoric is the more frequent the less complex the narrative structure is*. » (Boretzky 2000b : 44-45)

Tableau 13 : Fonctions de l'article défini selon (Boretzky 2000b : 44-54)

Type d'usage	Définition	Sous-type d'usage	Définition	Exemple	Traduction	Commentaire
1. Pragmatique	référence à la situation d'énonciation ou à des éléments préalablement mentionnés dans le discours	**anaphorique**	référence à un élément préalablement mentionné dans le discours	Ajia Varvara (Igla 1996 : 252) *… čhin iče duj-tane phabaja… alusardas e phabaja.*	« … coupe deux pommes… il choisit les pommes »	
		cataphorique	référence à un élément qui s'apprête à être défini (*via* une proposition relative, un adjectif, un syntagme génitival…)	Slovaquie (Ferková 1996 : 34) *o pharipen, so la sas andr-o jilo*	« la mélancolie qu'elle avait au cœur »	La référence à la mélancolie s'effectue grâce à la référence au cœur, dans la relative.
		situatif direct	référence à la situation concrète d'énonciation, quasi-déictique	Ajia Varvara (Igla 1996 : 253) *Kon mudardas e sapes?*	« Qui a tué le serpent ? »	La question est posée en présence du serpent.
2. Sémantique	référence à une connaissance partagée avec l'interlocuteur	**associatif anaphorique**	référence à un élément non encore mentionné, mais lié à un élément déjà mentionné	Ajia Varvara (Igla 1996 : 252) *… čhin iče duj-tane phabaja… e kožes da čho ka ko gras.*	« … coupe deux pommes… donne la peau à ton cheval. »	La référence à une pomme implique que celle-ci ait une peau (métonymie).
			référence à une situation qui comprend l'élément nouveau	Slovaquie (Boretzky 2000b : 49) *te džavas avri te bavinel pr-e dvoura le čhavenca*	« quand j'allais dans la cour pour jouer avec les enfants, … »	La référence aux enfants inclut l'idée de jeu, qui, par défaut, se fait dans la cour.
		usage situatif abstrait	référence à un élément faisant partie de la culture générale, des coutumes ou de l'expérience partagée	Gurbet de Bosnie (Boretzky 2000b : 51) *Vov lija e džukle thaj gelo*	« Il prit le chien et s'en alla. »	Le locuteur considère que la possession d'un chien est un fait attendu. Beaucoup de gens ont un chien et lorsqu'ils sortent, ils sont accompagnés de leur chien.
			référence à une entité unique supposée connue	*o kham* *o Devel*	« le soleil » « Dieu »	
			référence à une institution (école, organisme, instance de l'État) ou à une nation	Slovaquie (Boretzky 2000b : 51) *sar e Monika gel'a avri andal-e škola* Bugurdži (Boretzky 1993 : 163) *Sjomas momkos, džavas k-o gadžo.*	« quand Monika a terminé l'école » « J'étais jeune, j'allais régulièrement chez le Non-Rom [= chez le paysan serbe]. »	Il s'agit de l'employeur du locuteur, qui est considéré comme une institution : quelqu'un qui ne possède pas sa terre travaille forcément pour un patron.

Type d'usage	Définition	Sous-type d'usage	Définition	Exemple	Traduction	Commentaire
			référence temporelle (ponctuelle, durative, approximative)	Slovaquie (Boretzky 2000b : 52) *tel o trin kurke* Bugurdži (Boretzky 2000b : 52) *pal o duj zis* *M65A, Korçë, 1er août 2014* *to saati o duj*	« en trois semaines » « après deux jours » « à deux heures »	
		anamnéstique	référence à une connaissance spécifique partagée	Kalderaš de Voïvodine (Boretzky 1994° : 190) *bišalda ma munřo papo te des ma e sita*	« Mon grand-père m'a envoyé pour que tu me donnes le tamis. »	Le locuteur et l'interlocutrice savent que celle-ci possède un tamis.
3. Autre	–	**noms propres**	référence à un anthroponyme et à un toponyme	Slovaquie (Boretzky 2000b : 50) *e Praha* *e Marija*	Prague Marie	
		dans la comparaison	comparaison avec les noms indénombrables, abstraits, parfois dénombrables	Ajia Varvara (Igla 1996 : 85) *Severim tut sar o sekeri / o lon.*	« Je t'aime comme le sucre / le sel. »	Usage générique : référence aux propriétés d'une classe ou d'une partie d'une substance (Boretzky 2000b). Cet usage, peut-être un calque du grec ou du roumain, se trouve dans les variétés balkaniques.
		avec les noms indénombrables	compléments d'origine, compléments de matière	Bugurdži (Boretzky 1993 : 19) *katar o zlatos / i loš* Kalderaš de Voïvodine (Boretzky 1994° : 97) *katar o najlono*	« d'or » / « de joie » « en nylon »	Usage générique : référence aux propriétés d'une classe ou d'une partie d'une substance (Boretzky 2000b). Il pourrait s'agir d'une influence du grec.
		avec les noms dénombrables		Kalderaš de Voïvodine (Boretzky 1994a : 97) *porja katar e khanji*	« plumes de poule »	

Les usages pragmatiques et sémantiques semblent recouper la distinction entre anaphorique et non-anaphorique établie par Dryer (2013). Le premier usage (pragmatique chez Boretzky / anaphorique chez Dryer) fait référence à la situation d'énonciation, tandis que le second (sémantique chez Boretzky / non-anaphorique chez Dryer) fait référence à la connaissance partagée avec l'interlocuteur. Selon Matras (2002 : 98), les usages pragmatiques de l'article défini sont très classiques (entité identifiable parce que déjà donnée dans le discours ou accessible dans la situation d'énonciation), de même que les usages sémantiques (entité identifiable parce que présente dans la culture générale des interlocuteurs), à l'exception des usages situatif abstrait, anamnéstique et avec les noms propres, qui sont plus rares mais rappellent les autres langues balkaniques, dont le grec. L'origine déictique de l'article pourrait expliquer les usages pragmatiques et sémantiques : selon Matras (2002 : 112), l'article défini est un marqueur d'accessibilité contextuelle et figure entre l'ancien pronom personnel *lo* et le nouveau démonstratif *odova* sur l'échelle de continuité topique. C'est pourquoi il a pu calquer le but fonctionnel de l'article défini grec.

On constate que tous les usages de l'article défini sont spécifiques, qu'ils soient pragmatiques ou sémantiques. Les quatre « autres usages » sont à scinder en deux grands types : celui avec les noms propres est spécifique, tandis que les trois autres sont génériques, qui consistent à faire référence à un ensemble dont la référence est collective – ce sont les seuls usages génériques que peut revêtir l'article défini romani. Boretzky (2000b : 53) précise que l'on manque de données sur les variétés individuelles. On peut en effet se demander si ce n'est pas le type de corpus enregistré qui induit la rareté des occurrences de générique, ou bien si c'est le manque de diversité des types de corpus.

Le romani d'Albanie présente le même panel de fonctions. Les textes procéduraux du corpus présentent de nombreux usages sémantiques situatifs abstraits : la plupart des recettes commencent par « *On prend le x* ». L'exemple (12) constitue le début de la recette du bourek[1] (Figure 24) :

(12) *Romani (F71A, Bilisht, 27 juillet 2014)*

I	*marikli/*	*ka*	*l-es*	*o*	*vaɣo,*
ART.DEF.F.SG	bourek	AUX.FUT	prendre-2SG	ART.DEF.M.SG	farine(M.SG)
ka	*ćer-es*	*xumer,*	*ka*	*čhuv-es*	*æ̃ɣ-e.*
AUX.FUT	faire-2SG	pâte	AUX.FUT	mettre-2SG	œuf-M.PL

« Le bourek/ on prend la farine, on en fait de la pâte, on met des œufs. »

[1] Il s'agit d'une tourte populaire dans tout le bassin méditerranéen, à base de pâte feuilletée et farcie au fromage, à la viande, aux légumes ou aux pommes de terre.

Figure 23 : Locutrice de l'exemple (12)

Puisqu'il s'agit du début de la recette, aucun référent n'a encore été donné dans le discours. La recette est certes enregistrée près du feu (Figure 23), mais la préparation n'est pas visible. On constate que la pâte (*xumer*) et les œufs (*æ̃ɣe*) sont des noms nus, spécifiques mais indéfinis – tandis que la farine, elle aussi spécifique, est introduite par un article défini. Elle n'a pourtant été ni mentionnée, ni individualisée : la locutrice la considère un élément de la culture générale, de ma connaissance du monde. Je suis censée savoir que la marikli est à base de farine. Cet emploi n'est pas pragmatique (anaphorique) mais sémantique (non-anaphorique) : il est associatif puisque l'interprétation de bourek suppose farine : farine ⊆ bourek (Gaiffe, Reboul et Romary 1997). Huit des onze recettes du corpus présentent cet usage de l'article défini – les trois autres comportant un numéral ou un article zéro.

Figure 24 : Cuisson du bourek (*marikli*) au feu de bois à Bilisht

Boretzky (2000b : 53) précise que l'usage anamnéstique de l'article défini se différencie peu de l'usage situatif abstrait, à ceci près que l'on peut substituer un démonstratif au premier mais non au second. C'est le cas du romani d'Albanie :

(13) *Romani (M71A, Korçë, 28 juillet 2014)*

Čhuv-av-as	*odova*	*maja*	*jaurt-e-sele*
mettre-1SG-IMPF	DEM.F.SG	levure(F.SG)	yaourt-M.SG.OBL-(?)

« Je mettais cette levure dans le yaourt. »

Le locuteur a déjà mentionné la préparation du yaourt mais non la levure : « Moi il fallait/ que je traie les vaches/ que je fasse le lait/ que je prépare/ du yaourt/ que je fasse le fromage/ je faisais. […] Je fai/ je faisais le lait/ je préparais du yaourt/ yaourt tu connais ça le yaourt. Je mettais cette levure dans

le yaourt/ je faisais bouillir le lait/ je le laissais refroidir/ et je faisais du yaourt. » L'usage n'est donc pas pragmatique mais sémantique. Il fait référence à une connaissance partagée : le yaourt contient des bactéries lactiques qui le font fermenter.

1.3.1.6. Usage avec les noms propres

Les anthroponymes et toponymes doivent être accompagnés d'un article défini. C'est le cas pour toutes les variétés, même celles qui en ont perdu l'usage avec les noms communs, comme en Slovaquie ou en Pologne. Selon Boretzky (2000b : 53), les noms propres sont considérés comme uniques et ne peuvent être employés avec l'article indéfini : ne restent donc que l'article zéro et l'article défini, en fonction des variétés. Il semble s'agir d'une fonction particulière de l'article défini, ni pragmatique ni sémantique, mais en quelque sorte arbitraire puisque l'article est obligatoire sans apporter de nuance sémantique particulière. On trouve un tel usage dans d'autres langues des Balkans, comme le grec et l'albanais.

Boretzky (2000b : 60) note que l'article défini peut à l'occasion être omis en romani, notamment au cas oblique. Il donne pour exemple *e Praha*, Prague et *ø Kašate*, à Košice. Selon nous, cette omission est restreinte au cas des toponymes et à la fonction de complément de lieu, à sens statique :

(14) *Romani (M71A, Korçë, 28 juillet 2014)*

Akate	*si*	*o*	*džamutro/*	*Korča*
ici	être.3.PRES	ART.DEF.M.SG	gendre(M.SG)	Korçë

« Ici il y a le gendre/ à Korçë »

(15) *Romani (F47M, Llakatund, 13 août 2014)*

Akova	*beš-en*	*Elbasan.*	*ðeðe*	*me*	*isin-om*	*akate*	*Vloɻ.*
DEM.PL	habiter-3PL.PRES	Elbasan	et	1SG	être-1SG.PRES	ici	Vlorë

« Eux ils habitent à Elbasan et moi je suis ici à Vlorë. »

… ou dynamique (élatif ou directif) :

(16) *Romani (M71A, Korçë, 28 juillet 2014)*

Ka	*naš-en*	*Albanija-tar*	*ta*	*dž-an/*	*Amerika.*
AUX.FUT	fuir-3PL	Albanie-ABL	et	aller-3PL	Amérique

« Ils vont partir de l'Albanie et aller en Amérique. »

1.3.2. Article indéfini

L'indéfinitude est encodée au singulier par le numéral *(j)ek*, un, fléchi en genre et en cas, ou par l'« article » zéro. Le marquage de l'indéfinitude est selon Matras (2002 : 98) facultatif.

1.3.2.1. Jekh

Jek(h)[1] est le numéral cardinal « un ». Il a deux valeurs, celle de quantification et celle d'indéfinitude, d'où une ambiguïté fréquente en fonction du contexte. Le processus de grammaticalisation vers l'article indéfini, NUM > IND, semble en cours si ce n'est achevé dans de nombreuses variétés en romani. C'est le cas en romani des Balkans, comme d'ailleurs dans les autres langues de l'aire (Friedman 1999, Friedman 2000b, Friedman 2001, Friedman 2003a, Friedman 2003b). Les langues présentant le plus haut degré de grammaticalisation seraient le roumain, l'albanais et le turc, suivis du grec, du romani et du slave balkanique (Friedman 2001 : 292). L'auteur montre que le numéral est employé dans des contextes où la référence n'est pas définie, et qu'il est grammaticalisé au même titre que l'article défini, en romani comme dans les autres langues des Balkans (Friedman 2001 : 288). Étant donné qu'un tel « marqueur de l'indéfini » est quasiment inexistant dans les langues les plus anciennement attestées (grec ancien, vieux slavon d'église, latin), il s'agirait d'une innovation récente, à caractère aréal.

Jekh se fléchit comme un adjectif de la classe minoritaire, ne distinguant pas le genre au cas direct, mais distinguant éventuellement le genre et le nombre au cas oblique (strate 1). Notons que dans la variété de Parakalamos, il peut parfois se fléchir comme un nom, au moyen de la strate 2 (cf. section 1.3.5.1, p.127) : *jekhesa kaljardesa*, avec un soldat, au lieu de *jekhe kaljardesa* (Matras 2004 : 75), *i istoria isi njekeske mindeske dženeske*, l'histoire est à propos de l'un des miens, au lieu de *njeke minde dženeske.* La flexion de type nominal alterne cependant avec la flexion classique, de type adjectival pour *njek* : *o čikoro čhavo garádilo pala njeke rukhestar*, le petit garçon se cacha derrière un arbre, *ov mardas i kahnja njeke čhurjasa*, il tua la poule avec un couteau, ainsi que pour *jekh* : *oj phirela palal jeke muršestar*, elle marche derrière un homme. La flexion (de type nominal ou adjectival) semble elle-même optionnelle pour *jekh* : *dikhlem jek grastes to veš*, j'ai vu un cheval dans le bois, *kana šomas terni dživdilom jek bareste khereste ti poli*, quand j'étais jeune je vivais dans une grande maison en ville (Matras 2004 : 80).

Friedman (2001) compare les usages de l'article indéfini et constate qu'il s'agit essentiellement, en romani, d'usages spécifiques et chaque fois similaires à ceux de la langue de contact (bulgare pour l'arli de Bulgarie, serbe pour le gurbet de Serbie)[2]. Le numéral a un usage différencié selon les variétés. Il désigne une quantité indéfinie lorsqu'il précède un autre numéral en bugurdži : *jek deš zis*, une dizaine de jours, litt. « un dix jours » (Boretzky 1993 : 51). Dans ce cas, on peut dire que *jek* est employé à la fois comme numéral et comme indéfini. Selon Cech (2006 : 20), le parler de Novo Mesto et de Bela Krajina (Slovénie) utilise *jekh* de manière emphatique pour exprimer l'indéfini spécifique de sens « un certain » : *Do jekh veš kher hine, There was a house in some forest*. Il a également ce

[1] La consonne finale est aspirée ; si toutefois le mot suivant commence par une consonne, ou bien si *jekh* est suivi d'une pause, la consonne finale [k] n'est pas aspirée. J'ai remarqué, chez mes hôtes, que la désaspiration gagnait du terrain et que *jekh* tendait à être prononcé *jek* en toutes circonstances. Une variation dialectale touche en outre le jod initial (qui peut être amuï) et la voyelle ([ɛ] ou [ə]).

[2] Pour confirmer que l'usage de l'article indéfini du romani s'aligne sur celui des langues de contact, il faudrait cependant compléter cette étude par un travail sur corpus (semi-)spontané. Le romani étudié est un romani écrit, la traduction de la Bible, donc d'un texte sacré, probablement très littérale, présentant donc des calques.

sens de déterminant indéfini « un certain » dans le Burgenland (Halwachs et al. 2002 : 15). En kalderaš du Banat serbe, *jek(h)* est réduit à *ek, ək* voire simplement *k* lorsqu'il est employé en tant qu'article indéfini, ce qui n'est pas le cas du numéral, toujours prononcé *jek(h)* (Boretzky 1994a : 58, 94). On peut parler d'une scission paradigmatique entre *jek* numéral, et *(j)(e/ə)k* article indéfini : dans cette variété, la grammaticalisation de l'article indéfini est accomplie.

Quelle que soit sa fonction, *jekh* introduit toujours un référent spécifique, mais jamais générique (Boretzky 1993 : 21, Boretzky 1994a : 98), contrairement à l'article défini qui peut assumer un sens générique. L'article indéfini revêt selon Matras (2002 : 98) les fonctions suivantes :

1) introduire dans le discours un nouvel élément indéfini topique

Matras (2002 : 98) emprunte à Boretzky (1994a : 229) l'exemple suivant : *sas haj sas jek gaźo čoŕo*, 'there was a poor man' (début d'un conte). On trouve cette fonction en arli d'Albanie :

(17) *Romani (M46A, Korçë, 22 juillet 2014)*

Phen-l-as	*ə:*	*man-ge*	*mi*	*mami*	*kaj*	*ə:* […]
dire-PST-3SG	euh	1SG.OBL-DAT	POSS.1SG-F.SG	grand-mère(F.SG)	COMPL	euh
kaj	*jek*	*phur-i*	*em*	*jek*	*ə*	*jek*
COMPL	IND	vieux-F.SG	et	IND	euh	IND
phur-o	*beš-ena*	*isine*	*korkorik*	*jek*	*ćher-es-te.*	
vieux-M.MG	rester-3PL	AUX.IMPF	seul	IND	maison-M.SG.OBL-LOC	

« Ma grand-mère euh m'a raconté que euh […] qu'une vieille et un euh un vieux habitaient seuls dans une maison. »

2) discriminer un élément indéfini au sein d'un tout (lieux, événements)

Matras (2002 : 98) emprunte à Boretzky (1994a : 236) l'exemple suivant : *sas duj phral and ek foro, 'there were two brothers in a [particular] town'*. C'est le cas de *jek*[1] *ćhereste*, dans une maison (particulière) dans l'exemple (17). *Jekh* est également un pronom (18) et un adjectif (19) :

(18) *Romani (M22A, Korçë, 6 août 2014)*

Jek	*isi*	*po*	*bar-i*	*man-da*	*jek*	*isi*	*po*	*tikn-i.*
IND	être.3.PRES	COMP	grand-F.SG	1sg.OBL-ABL	IND	être.3.PRES	COMP	petit-F.SG

« L'une (litt. une) est plus grande que moi l'autre (litt. une) est plus petite. »

(19) *Romani (F60A, Korçë, 21 juillet 2014)*

Jek	*isi*	*amar-i*	*čhib*	*ə/*	
IND	être.3.PRES	POSS.1PL-F.SG	langue(F.SG)	euh	
sa	*i*	*roman-i*	*čhib*	*jek*	*isi.*

[1] Notons qu'ici la locutrice ne fléchit pas *jek* au cas oblique (on attendrait *jekhe ćhereste*).

tout	ART.DEF.F.SG	romani-F.SG	langue(F.SG)	IND	être.3.PRES

« Notre langue est une euh/ toute la langue romani est une seule et une seule (litt. une). »

La fonction de l'adjectif peut être attribut du sujet (19), mais aussi épithète du nom. Si le syntagme est introduit par un article défini, on a la séquence *o jekh phral*, *one of the brothers*, l'un des frères, littéralement « le un frère » (lovari, Gjerde et Kristiansen 1994 : 24, cité par Matras 2002 : 166). Un tel usage révèle l'ambivalence du terme entre numéral et indéfini. Une telle séquence est d'ailleurs possible en grec moderne (exemple (77).[1] Elle se rencontre en arli d'Albanie :

(20) *Romani (M46A, Korçë, 24 juillet 2014)*

O	*jek*	*tikn-o*	*phral*	*isi/*	*vaćer-en*		*les-će*
ART.DEF.M.SG	IND	petit-M.SG	frère(M.SG)	être.3.PRES	parler-3PL.PRES		3SG.OBL-DAT
o	*duj*	*phral-a*	*bar-e/*	*eh*	*akowa*	*isi*	*delin-o.*
ART.DEF.M.SG	deux	frère-M.PL	grand-PL	eh	DEM.M.SG	être.3.PRES	fou-M.SG

« L'un des frères, le petit (litt. le un petit frère) est/ ils lui disent les deux grands frères/ eh mais il est fou lui. »

1.3.2.2. Article zéro

Dans les autres contextes indéfinis, le nom n'est accompagné par aucun déterminant.

1) Selon Matras (2002 : 98), c'est le cas des **entités indéfinies non topicales, ne pouvant pas être ambiguës et ayant une fonction « ancrante » ou d'attribut du sujet**. Il donne un exemple suivant en lovari : *Traisardam kothe, sas ame kher, muro dad puterdas kirčima*, 'We lived there, we had [a] house, my father opened [a] pub'. En romani de Parakalamos : *oj bičhavdas mange lil*, elle m'envoya une lettre (Matras 2004 : 80). Selon Adamou (2016 : 158), les constructions attributives à copule être/devenir ou sans copule sont suivies de nom nus en romani de Thrace. On trouve ce contexte en arli :

(21) *Romani (F16AČ, Korçë, 22 juillet 2014)*

O	*ćermuso*	*bar-il-o,*	*ondil-o*	*čhavo*	*bu:t*	*but*	*lač-o*
DEF.M.SG	souris(M.SG)	grand-PST-M.SG	devenir.PST-M.SG	garçon(M.SG)	très	très	beau-M.SG

« La souris grandit et devint [un] trèèès très beau garçon. »

Le référent de ces deux exemples est spécifique. Dans l'exemple suivant au contraire, il est quasiment générique : *Thodem ma t'avav sluga, Ich ließ mich als Knecht anstellen*, Je me fis serviteur (Boretzky 1994a : 99). Dans ce type de contexte, il existe un flottement entre article indéfini et nom nu selon

[1] Plank (2003b : 369-371) étudie ce type de constructions, que l'on trouve entre autres en moyen-haut allemand, en italien qu'il nomme « pré-standard » et en anglais.

(Boretzky 1994a : 99) : *Sas ma (k) źukəl, Ich hatte [einen] Hund*, J'avais [un] chien ; *Te arakhlemas mangə stano, daß ich mir [eine] Wohnung gesucht hätte,* que je me serais cherché [un] appartement.

2) Lorsqu'**un objet est nié dans une tournure possessive**, il ne porte pas d'article : *Naj ma khər, Ich habe kein Haus*, Je n'ai pas de maison (Boretzky 1994a : 99).

3) Lorsqu'**un objet non défini est un nom de matière indénombrable**, il ne porte pas d'article : c'est le cas de *xumer*, pâte à tarte, dans l'exemple (12). Boretzky (1994a : 99) parle de « matériaux non individualisables » mais précise que s'ils font référence à une entité déjà mentionnée, ils peuvent être accompagnés d'un article défini. Il oppose ces deux exemples :

(22) Romani (Boretzky 1994a:99)

Pher-el	*pesk-o*	*brek*	*praxo*
remplir-3SG.PRES	POSS.REFL.SG-M.SG	poitrine(M.SG)	cendre

Er füllt seine Brust (unter dem Hemd) mit Staub. (« Il remplit sa poitrine de cendre. »)

(23) Romani (Boretzky 1994a:99)

Trago	*katar*	*o*	*praxo*
trace	venant_de	ART.DEF.M.SG	cendre(M.SG)

eine Spur von dem (besagten) Staub (« une trace de la cendre [déjà mentionnée] »)

L'usage des noms non-individualisables avec article zéro s'étend selon Boretzky (1994a : 99) aux **objets individualisables**, dès lors que le référent n'est pas spécifique ou que l'on peut se représenter comme un matériau : *Peklja khajni, Sie briet (ein) Huhn*, Elle fit cuire une poule. On trouve cependant des exemples avec article zéro où le référent est spécifique : *Von dikhən baro paj, Sie sehen ein großes Wasser (im Sinne von Gewässer),* Ils voient une grande étendue d'eau (litt. grande eau). Pour cette expression, on peut hésiter entre interprétation dénombrable et indénombrable, mais elle est spécifique : c'est pourquoi on attendrait plutôt l'emploi de *jek*.

4) Lorsqu'un **nom non défini a un sens abstrait**, il ne porte pas d'article. C'est le cas des concepts abstraits, que le romani traite comme des indénombrables. Boretzky (1994a : 99) fournit un exemple en kalderaš : *Tunča sas baro čořipe, Damals war (eine) große Armut,* A l'époque régnait [une] grande pauvreté. On trouve ce contexte en arli :

(24) *Romani (F16AČ, Korçë, juillet 2013)*

Akana	*isi*	*nev-o*	*mentalitèti.*
maintenant	être.3.PRES	nouveau-M.SG	mentalité(M.SG)

«Maintenant il y a [une] nouvelle mentalité. »

Certains **phénomènes naturels indénombrables** se passent également d'article défini dans une tournure causale (Boretzky 1994a : 98) – comme en allemand et en français : *Merel bokhatar / rovel trušatar, Er stirbt vor Hunger / er weint vor Durst,* Il meurt de faim / il pleure de soif.

5) Les noms **dénombrables au pluriel** ne prennent pas d'article indéfini. C'est le cas d'*æ̃ɽe*, œufs, dans l'exemple (12). Il n'existe **pas d'article indéfini pluriel** dans la plupart des variétés. Certaines en ont créé un par contact. C'est le cas du kalderaš du Banat qui emprunte au roumain *nişte/neşte* son article *nešte*, des, quelques, certains (Boretzky 1994a : 31) : *Maladilem nešte ŕomenca, Ich traf mich mit einigen Zigeunern*, J'ai rencontré quelques Roms. Il peut également être utilisé au singulier : *Avel nešte manuš, Es kommt ein (bestimmter) Mensch,* Arrive un (certain) homme (Boretzky 1994a : 100). Au pluriel, zéro et *nešte* sont équivalents – tandis qu'au singulier, zéro et *jek* sont souvent en distribution complémentaire.

6) Les **noms propres au vocatif** ne portent pas d'article (Adamou 2016 : 158).

7) Certaines **expressions où est répété un nom non défini** ne présentent pas d'article : *gjes pala gjeste*, *Tag nach Tag (eines (schönen) Tages)*, un beau jour (litt. jour après jour) (Boretzky 1994a : 99). C'est également le cas des énumérations.

8) Les **locutions à verbe support** présentent le plus souvent un article zéro. Ces constructions combinent un verbe de sens général ou abstrait à un nom portant l'essentiel du sens : ainsi en français *avoir faim/peur/honte/mal/raison/tort/besoin…*, *faire honneur/honte/mal/pipi/ outrage/semblant/mine...*, *prendre congé/garde/racine…* On trouve dans les variétés d'Albanie les expressions suivantes : *kerav buti*, je travaille (litt. je fais travail), *kerav lafi*, je parle (litt. je fais parole), *dav vast*, j'aide (litt. je donne main), *del iv*, il neige (litt. il donne neige), *xav maro*, je mange (litt. je mange pain), *xal daba*, je reçois des coups (litt. je mange coups). Cette approche nous semble expliquer certains cas difficiles de Boretzky (1993 : 21), comme *ka keras biav, wir werden Hochzeit feiern*, nous allons fêter un mariage.

L'exemple arli suivant synthétise trois contextes d'indéfinitude. *Manuš,* être humain, est un nouvel élément indéfini topique, d'où l'usage de l'article indéfini *jek*. *Godi,* pensée, est un nom indénombrable, d'où l'absence d'article. *Lafi,* parole, et *lira,* lire, sont des noms dénombrables comptabilisés pour exprimer un échange, d'où l'usage du numéral *jek,* un.

(25) *Romani (M71A, Korçë, 28 juillet 2014)*

Isine	*jek*	*manuš*	*kaj*	*bikin-ela*	*godi!*	*jek*	*lafi,*	*jek*	*lira.*
être.3.PRET	IND	être_humain	REL	vendre-3SG	pensée	IND	parole	IND	lire

« Il y avait un homme qui vendait de la pensée ! une parole, une lire. »

1.3.3. Démonstratif

1.3.3.1. Formes et emplois

Le romani possède un système de démonstratifs à quatre termes qui encode à la fois la présence immédiate ou l'absence du référent, et le degré de spécificité ou d'ambiguïté de celui-ci (Matras 2011 : 263-264). Ce système serait hérité du romani ancien (Matras 2004 : 78). Les formes sont extrêmement variables en fonction des dialectes mais présentent une structure commune (Matras 2002 : 103-105). Le mečkar et l'arli d'Albanie présentent les formes[1] du Tableau 14 :

Tableau 14 : Formes du démonstratif en mečkar et arli

		non-spécifique		spécifique	
		situation	**discours**	**situation**	**discours**
M.SG.DIR	**mečkar & arli**	*akova*	*odova*	*akava*	*okova*
F.SG.DIR	**mečkar & arli**	*akoja / akəja*	*odoja / odəja*	*akaja / akəja*	*okoja ([okəja])*
PL.DIR	**mečkar**	*akola*	*[odola]*[2]	*akala*	*[okola]*[3]
	arli	*[akole]*	*odole*	*akale*	*okole*
SG.OBL	**arli**	*odole*	(?)	*akale*	*okole*
PL.OBL		(?)	(?)	*akale*	(?)

Les deux variétés ne se distinguent qu'au pluriel : le mečkar a conservé la désinence *-a* tandis que l'arli l'a remplacée par *-e*, probablement par analogie avec la désinence adjectivale. Parmi les formes de singulier, on note que la voyelle radicale est le plus souvent centralisée au féminin. Les formes spécifique et non-spécifique ne sont plus distinctes : *akoja* et *akaja* sont confondues en *akëja*. Ce phénomène s'étend à l'occasion au masculin : le corpus présente une occurrence de *akəva*. Les formes obliques se rencontrent moins fréquemment. À l'oblique en arli, la différence de genre n'est pas maintenue et les formes sont les mêmes qu'au direct pluriel. Matras (2002 : 105) note que le paradigme du romani ancien (Matras 2002 : 110) est conservé tel quel dans les variétés périphériques : extrême-occidentales (Pays de Galles) et du sud-est, dont l'arli du Kosovo et de Macédoine. On peut y ajouter l'arli et le mečkar d'Albanie, à l'exception notable de la série non-spécifique situationnelle *adava* et peut-être du paradigme de l'oblique.

1.3.3.2. Distribution et double détermination

Le démonstratif, en tant que déterminant, occupe canoniquement la première place dans le syntagme nominal (Matras 2002 : 105, 166). (Boretzky (1993 : 48) a cependant constaté qu'en bugurdži, le démonstratif peut être postposé, et dans ce cas « pleinement fléchi », comme un nom aux strates 1 et

[1] Les formes entre crochets ne sont pas attestées dans notre corpus ; nous les postulons ici pour compléter le paradigme.
[2] *Odola* n'est pas attesté dans le corpus mečkar et arli, mais il l'est dans le corpus čergar.
[3] *Okola* n'est pas attesté dans le corpus mečkar et arli, mais il l'est dans le corpus džambazo (Macédoine).

2 *: e rakleske kodleske, jenem Jungen,* ce garçon non-rom. Le démonstratif est dans cet exemple non seulement postposé, mais également accompagné d'un article défini. Il y a alors double détermination (DEF N DEM) (cf. section 1.3.3.2, p. 122). En kalderaš du Banat serbe, le démonstratif peut être postposé en cas d'emphase (Boretzky 1994a : 55) :

> « *Als Attribute gehen die Dem. dem Nomen voran, bei Emphase können sie auch die zweite Stelle einnehmen; vgl.* kîlcî kodola*. In dem Fall werden eher die voll ausflektierten Formen gebraucht ; vgl.* e balen kodolen *„jene Schweine",* e gaźeskə kodoleskə *„jenem Mann" und selbst vorangestellt* kodoles maśes *„jenen Fisch". In der Kombination Pron. + Subst. wird normalerweise der Artikel nicht gesetzt, also z.B.* kaja sobica *„dieses Zimmerchen", aber der Artikel ist zuminstest in S nicht ganz selten :* kado kaš o baro *„dieser große Baum" bzw. „dieser Baum, der große".* » (Boretzky 1994a : 55)

On a ici un exemple de postposition simple du démonstratif (N DEM) et deux exemples de postposition du démonstratif avec présence de l'article défini (DEF N DEM), ainsi que deux exemples d'antéposition simple du démonstratif (DEM N) et un exemple d'antéposition du démonstratif avec présence de l'article défini (*kado kaš o baro,* DEM N DEF A). L'article défini, en position postnominale, semble accompagner l'adjectif plutôt que le nom : il y a non seulement double détermination, mais également polydéfinitude. L'auteur remarque que le démonstratif, lorsqu'il est postposé, est fléchi « pleinement », c'est-à-dire suivant les strates 1 et 2.

Le démonstratif se postpose et s'accompagne d'un article défini dans d'autres variétés des Balkans, dont celle d'Ajia Varvara (Igla 1996 : 40, 85). Cette combinaison n'existe pas hors des Balkans selon Matras (2002b : 96), et il s'agit d'un balkanisme tardif. L'article défini est facultatif lorsque le démonstratif est antéposé, et obligatoire lorsque le démonstratif ou le possessif sont postposés : « *Sas ekh thagár. Kadaá (o) thagár xurdé ninái les, Es war ein König. Dieser (der) König hatte keine Kinder* », Il était une fois un roi. Ce roi (litt. ce (le) roi) n'avait pas d'enfants, « *Mulás léski džuvlí /i džuvlí léski, Seine Frau / die Frau seine starb* », Sa femme (litt. la femme sienne) mourut. La double détermination est selon Igla (1996 : 85) pléonastique car l'article n'a pas d'apport sémantique ou fonctionnel particulier. On peut dès lors supposer qu'il sert à remplir la position initiale du déterminant, laissée vide, et qu'un tel article est en quelque sorte explétif.

La variété de Parakalamos offre une syntaxe intermédiaire : le démonstratif est habituellement antéposé au nom (Matras 2004 : 97). Mais il peut aussi être accompagné d'un article défini, auquel il est alors antéposé. On trouve donc des structures DEM DEF N, voire DEM DEF NUM DEF N : « ova o čhavo našadah but love *'that boy lost a lot of money',* kamesa te kheles olenza i trinenza i romaceljenza? *'do you want to play with those three Gypsy girls?'* » (cf. exemple (31). Ceci n'empêche pas de trouver également la distribution héritée du romani, DEM N : *ova kher,* cette maison. Lorsqu'il y a usage pléonastique des strates 2 et 3 pour exprimer une relation (cf. section 1.2.2, p. 106), le démonstratif est antéposé à l'ensemble du syntagme nominal, et même antéposé à la préposition (DEM PREP.DEF N) : *kapjos dživel aleste to kher*, quelqu'un vit dans cette maison (Matras 2004 : 79), *štar romacel dživena oleste to kher,* quatre Roms vivent dans cette maison (Matras 2004 : 97). Cette syntaxe va à l'encontre de l'ordre des constituants canoniques, et constitue une particularité de cette variété, due à l'influence du grec.

Le démonstratif occupe selon Matras (2000b : 103) une position « externe » au syntagme nominal, tandis que les adjectifs sont ajoutés de manière « interne » : DEM A (A) (A) … N. Occupant dans ce cas la place du determinant, « *Unlike adjectival attributes, the attributive demonstrative has intrinsic determiner quality which makes it generally incompatible with the definite article* » (Matras 2000b : 102). Citant Boretzky (1994a : 55) et Igla (1996 : 40, 165), il constate pourtant qu'il y a compatibilité dans certains dialectes et parle d'« *article integration in demonstrative determination* » lorsque les deux déterminants co-occurrent (Matras 2000b : 102), et notamment lorsque le démonstratif est postposé (Matras 2000b : 103). Matras (2000b : 102) propose une explication théorique à la morphologie du démonstratif postposé :

> « *It seems that the highly marked position of the demonstrative in (7)*[1] *indicates an afterthought to the core noun phrase, or at least a postponed modification, in the mental planning of the utterance as well as in linear ordering. The demonstrative thus assumes the status of a pronominal element, potentially independent of the head noun, i.e. capable of autonomous representation of a referent, which accounts for its attracting full or 'strong' case inflection.* » (Matras 2000b : 102)

L'auteur parle plus loin de « statut affaibli de determinant » au profit d'une « autonomie référentielle plus forte » (Matras 2000b : 103). Un démonstratif postposé porterait la « flexion forte » et serait accompagné d'un article défini en raison de son statut intermédiaire entre le déterminant et le pronom : « *We may thus speak of a continuum of demonstrative functions, from more pronominal to more attributive, with the postposed demonstrative figuring somewhere in-between* » (Matras 2000b : 103). L'emplacement postnominal d'un modifieur, selon lui hautement marqué, est qualifié d'*afterthought*, commentaire postérieur non prévu cognitivement, « ajouté après coup ». Matras (2002 : 105) reprendra l'hypothèse boretzkienne de l'emphase : « *Emphatic use of postposed demonstratives is quite frequent in some dialects, though here demonstratives carry nominal inflection and are compatible with definite articles, and so they may be regarded as appositions (*e romeske kodoleske *'ART man.DAT that.DAT = for that man'; cf.* kodole romeske*).* » L'auteur emploie ici le terme d'apposition, qui peut correspondre, en syntaxe, à la notion fonctionnelle d'*afterthought*.

L'origine de ces structures est selon Boretzky (1994a : 55) à chercher dans le contact de langues. Les constructions du kalderaš rappellent le roumain *acest om*, cet homme (DEM N) et *om-ul acesta*, cet homme (N-DEF DEM) où le démonstratif est postposé et le nom accompagné d'un article défini – même si l'ordre des mots est différent. La postposition du démonstratif sans article défini (N DEM), possible en romani, n'est cependant pas mentionnée pour le roumain. Igla (1996 : 165) propose elle aussi que le romani ait emprunté ces structures, cette fois au grec :

> « *Für das attributivische Demonstrativum bestehen die folgenden Anordnungsmöglichkeiten:*
>
> - *DEM SUBST* kadavá rom/kadalé romés
> - *DEM ART SUBST* kadavá o rom/kadalá e romés *„dieser/diesen Mann"*
> - *ART SUBST DEM* o rom kadavá/e romés kadalá
>
> *Die Typen 2) und 3) entsprechen dem Griechischen, vgl.* αὐτός ὁ ἄντρας *und* ὁ αὐτός ἄντρας *mit obligatorischer Setzung des bestimmten Artikels gegen das dem ersten Typ entsprechende* *αὐτός ἄντρας. *Daß die Verbindung des Demonstrativums mit dem Artikel unter rezentem griechischen Einfluß entstanden ist, läßt sich anhand einer Frequenzanalyse des Gebrauchs von Konstruktionen*

[1] *e gaźeskə kodoleskə*, ce Non-Rom, exemple tiré de Boretzky (1994a : 55).

mit und ohne Artikel bei älteren und jüngeren Sprechern (s. 5.5.4) wie auch aus dem Vergleich mit außerhalb Griechenlands gesprochenen Romani-Dialekten nachweisen. […] » (Igla 1996 : 165)

Selon Boretzky (2000b : 43-44), le type 1 (DEM N) est le type hérité ; le type 2 (DEM DEF N) est un type balkanique emprunté au grec, présent en kalderaš (Boretzky 1994a), à Ajia Varvara (Igla 1996) et chez les Tchinghianés ottomans (Paspati 1870) ; le type 3 (DEF N DEM) est emprunté au grec mais converge avec le roumain, présent en kalderaš et en ursaritiko (Boretzky 1994a). Un type 4 n'est pas mentionné, qui serait celui de Parakalamos, DEM DEF N – emprunté au grec selon Matras (2004 : 78). Matras (2002 : 167) ira plus loin en qualifiant la postposition du démonstratif de « calque », ce qui explique pourquoi il fait alors partie de l'emplacement postnominal qualifié d'« option » (section 1.4.3, p. 138). La postposition est selon lui « formelle » car elle ne relève pas d'une intention communicative ou d'une visée pragmatique. Selon Matras (2000b : 103), le romani d'Ajia Varvara et le kalderaš montrent que le romani a poursuivi la « balkanisation » de sa typologie syntaxique, en ce qui concerne les règles de détermination et l'ordre des constituants nominaux. Le fait, pour un démonstratif, d'être accompagné d'un article défini (DEM DEF N ou DEF N DEM) est un trait aréal puisqu'il se trouve en grec, en roumain, en romani *vlax* et balkanique (de Grèce, de Yougoslavie et de Roumanie), mais non dans les variétés hors des Balkans (Matras 2002 : 97).

1.3.4. Possessif

1.3.4.1. Formes

Les possessifs de première et deuxième personne du singulier et du pluriel présentent la même structure et sont de formation récente (Matras 2002 : 99, Elšík 2000). Les possessifs de troisième personne singulier, pluriel et réfléchi sont encore plus récents ; ils sont homophones des pronoms personnels déclinés au génitif. Le possessif peut présenter une forme longue et une forme courte aux trois personnes du singulier ainsi qu'au réfléchi (Boretzky 1993 : 46). La forme courte ne présente que le radical de personne et la désinence finale : *mo* au lieu de *mi(nd)řo/ mu(nd)řo*, *to* au lieu de *ti(n)ro* (Matras 2002 : 100). On observe une variation dialectale pour la forme courte :

- en bugurdži, elle est plus courante pour les première et deuxième personnes du singulier que pour la troisième personne et le réfléchi,
- à Ajia Varvara, elle n'existe pas à la troisième personne du singulier (Igla 1996 : 165),
- en kalderaš, elle n'existe que pour POSS.2SG (Boretzky 1994a : 54).

Lorsqu'il occupe la fonction de pronom, le possessif ne connaît au contraire que la forme longue et suit la flexion nominale (Boretzky 1993 : 47, Boretzky 1994a : 54). Notons que dans la variété de Parakalamos, la forme longue suit la flexion nominale même lorsqu'elle occupe la fonction de déterminant » (Matras 2004 : 76). Les variétés d'Albanie présentent les formes du Tableau 15 :

Tableau 15 : Formes du possessif dans les variétés d'Albanie

POSSESSIF			
1 SG	*m- / mir-* [*murn-* en čergar]	3 SG.M	*lesk-(r-)*
2 SG	*t- / tir-*	3 SG.F	*lak-(r-)*
1 PL	*amar-*	3 PL	*leng-(r-)*
2 PL	*tumar-*	REFL	*pesk-(r-) / pir-*

1.3.4.2. Distribution et double détermination

La distribution du possessif présente les mêmes caractéristiques que celle du démonstratif, du moins dans les variétés des Balkans et *vlax* (Boretzky 2000b : 44). Selon Boretzky (1993 : 19-20, 46-47), le possessif est antéposé au nom mais peut également lui être postposé si un contraste doit être établi entre deux possesseurs – ce qui n'est pas sans rappeler l'idée d'*emphase* évoquée pour la postposition du démonstratif. Selon Soravia (1972 : 38), le possessif serait postposé au nom en romani des Abruzzes en raison de l'influence de l'italien familier. Igla (1996 : 165) précise qu'à Ajia Varvara, la forme courte de POSS.1SG, POSS.2SG et POSS.REFL peut être seulement antéposée, mais que la forme longue, « plus marquée », peut être soit anté- soit postposée. Les possessifs des autres personnes (3SG, 1PL, 2PL, 3PL) peuvent indifféremment être anté- ou postposés sans effet d'emphase. L'usage d'un article défini précédant le nom est obligatoire lorsque le possessif est postposé, plus encore que dans le cas du démonstratif selon Boretzky (1993 : 19-20) et Igla (1996 : 165). La distribution du possessif est donc POSS N (*mi/miní dej*, ma mère) et DEF N POSS (*i dej miní*, ma mère) (Igla 1996 : 165). Il existe toutefois des asymétries entre la syntaxe de ces deux déterminants :

- L'antéposition (canonique) reste plus fréquente pour le possessif que pour le démonstratif. C'est pourquoi la double détermination avec DEF est également moins fréquente (Boretzky 2000b : 43). Son existence s'explique également par le contact de langues, plus exactement par l'influence du roumain sur les variétés *vlax* : *maic-a mea*, ma mère (N-DEF POSS) est l'ordre roumain non-marqué par rapport à *mea maic-a*, ma mère (POSS N-DEF). Il est selon nous difficile d'expliquer pourquoi l'influence du roumain aurait été moindre sur la syntaxe du possessif que sur celle du démonstratif.
- Il n'existe pas de structure POSS DEF N **mi/miní i dej*, ma mère, parallèle à DEM DEF N *kadavá o rom*, cet homme. Cette syntaxe existe pourtant dans les autres langues, par exemple bulgare et macédonien (« moja-ta kniga,*"my-the book"* », mon livre, Boretzky 2000b : 43).
- L'existence de deux formes permet en revanche au possessif d'être dupliqué dans certaines variétés : en tchinghiané ottoman « minri mi dukani *litteraly "mý my darling",* tinre te vast *"yóur your hands"* » (Boretzky 2000b : 43-44) ; en arli de Prilep *mindo mo dat*, mon père, *mindi mi daj*, ma mère (Heinschink 1978 : 14) ; en romani de Thrace « mirne me vasta *'lit. my hands of mine',* mirni mi dej *'lit. my mother of mine'* » où ces constructions sont très rares (Adamou 2016 : 161) ; à Parakalamos *sigo ka me therav but love ka te kinav* <u>*mindo mo*</u>

aftokindos, bientôt moi j'aurai assez d'argent pour acheter ma propre voiture, où cette construction permet de mettre l'accent sur la propriété exclusive (Matras 2004 : 76).

- Certaines variétés dupliquent le possessif lorsque le nom est accompagné d'un adjectif ou d'un numéral, ainsi à Parakalamos (Matras 2004 : 76) : *me duj me phenja*, mes deux sœurs (POSS NUM POSS N : cf. exemple (32), *mo baro mo phral*, mon grand frère (POSS A POSS N).

Le possessif peut enfin se combiner à d'autres déterminants que l'article défini, à commencer par le démonstratif (Igla 1996 : 165-166) :

« *DEM POSS SUBST*	kodová lésko phral
DEM ART SUBST POSS	kodová o phral lésko *„jener sein Bruder"*
POSS SUBST DEM	lésko phral kodová »

On note que le deuxième exemple (DEM DEF.ART N POSS) est un cas de triple détermination. Enfin, le possessif peut se combiner à *jek* dans « *još jek mnřo pobrátimo, noch einer meiner Kumpel* », encore l'un de mes camarades/frères (Boretzky 1994a : 54), où celui-ci joue un rôle de numéral. On trouve cette combinaison dans la variété de Parakalamos : « i istoria isi njekeske mindeske dženeske *'the story is about a relative of mine'* » (Matras 2004 : 76). Elle n'est pas sans rappeler *o jekh phral*, l'un des frères, et l'exemple (20), où *jekh* joue un rôle d'adjectif.

1.3.5. Quantifieur et numéral cardinal

On appelle quantifieur un élément donnant une information de quantité sur un autre élément, notamment un nom ou un verbe. Il s'agit d'une catégorie de mot plutôt sémantique, qui comprend les numéraux cardinaux (il y a controverse sur ce point), des quantifieurs dénombrables et indénombrables, de quantité nulle, singulière, plurielle, totale, collective et distributive. Certains quantifieurs sont dits universels car ils s'appliquent à l'ensemble des constituants du syntagme nominal. Du point de vue morphologique en revanche, il n'y a pas d'unité puisque certains quantifieurs sont variables, d'autres invariables. Du point de vue syntaxique, ils se placent avant ou après l'élément qu'ils déterminent. Mais les quantifieurs nominaux tendent à être antéposés au nom – qu'ils occupent une position et une fonction de déterminant ou d'adjectif.

1.3.5.1. Formes

Le principal quantifieur du romani est *sa/sja*, tout ; il est invariable et toujours antéposé au nom. L'adverbe *but*, beaucoup, peut jouer lui aussi cette fonction de déterminant quantifieur ; il est lui aussi invariable et antéposé au nom. On trouve également des quantifieurs empruntés, tels que les distributifs *s(v)ako* (slave) et *(h)er*, chaque (turc). Dans la variété de Parakalamos, le quantifieur peut être accompagné d'un article défini, auquel il est alors antéposé : *sare o kherimata si parne*, toutes les maisons sont blanches (Matras 2004 : 97).

Le romani présente un système numéral en partie hérité et en partie emprunté au grec (7, 8, 9, 30, 40) et aux langues de contact récentes. Le romani de Parakalamos a emprunté toutes ses dizaines et mille au grec, langue de contact ancienne, actuelle et récente (Matras 2004 : 75). L'arli et le mečkar d'Albanie ont seulement emprunté au grec 50 (*penda*, mais *peninda* chez un locuteur) voire 60 (*eksinda* chez certains locuteurs), et au slave 1000 (*xiljada*). Plus les nombres sont complexes et importants (pour exprimer un âge avancé ou un prix élevé), plus les locuteurs préfèrent les exprimer dans la langue de contact. C'est le cas en Albanie, mais également en Voïvodine, où les Kalderaš ont recours au serbo-croate pour les hauts numéraux cardinaux et ordinaux (Boretzky 1994a : 58). J'ai constaté que plusieurs informateurs mečkars ne connaissaient pas les noms de dizaines hérités (20, 60, 70, 80, 90) ni 100, et les remplaçaient par leurs équivalents albanais ou grec. D'autres enfin comptaient entièrement en albanais, à l'exception des premières unités. En arli, les nombres hérités terminant par une consonne ou une semi-consonne se fléchissent comme des adjectifs (exemple (26), notamment pour indiquer l'âge (exemple (27) – de même en romani dolenjski (Cech et Heinschink 2001 : 159) et en romani de Parakalamos (Matras 2004 : 75). Il arrive cependant que le numéral ne soit pas fléchi, à l'oblique (Matras 2004 : 75) et dans les expressions temporelles (Matras 2004 : 93).

(26) *Romani (M46A, Korçë, 24 juillet 2014)*

Odol-e	*len-ja-te/*	*ə/*	*odol-e/*	*bare*	
DEM-F.SG.OBL	rivière-F.SG.OBL-LOC	euh	DEM.F.SG.OBL	grand-F.SG.OBL	
len-ja-te	*isine*	*i*	*kučedra/*	*dešuduj-e*	*šer-en-sar.*
rivière-F.SG.OBL-LOC	être.3.PRET	ART.DEF.F.SG	dragon(F.SG)	douze-PL.OBL	tête-M.PL.OBL-INSTR

« Dans cette rivière/ euh/ cette/ grande rivière il y avait le dragon/ à douze têtes. »

(27) *Romani (F65A, Korçë, 1er août 2014)*

I	*čhaj*	*o-ndil-i*	*duj-e*	*breš-en-gir-i.*
ART.DEF.F.SG	fille(F.SG)	devenir-PRET-3SG.F	deux-OBL.PL	année-M.PL.OBL-GEN-F.SG

« La fille atteignit l'âge de deux ans. »

1.3.5.2. Distribution et double détermination

Les numéraux cardinaux, lorsqu'ils jouent le rôle de déterminant, occupent la première position, canonique, du syntagme nominal :

(28) *Romani (M58MČ, Voskopojë, 15 juillet 2013)*

Amen	*sin-am…*	*ʃtar*	*fis-ja*	*bar-e.*
1PL	être-1PL	quatre	peuple-M.PL	grand-M.PL

« Nous sommes… quatre grands peuples. »

Notons que le numéral peut être accompagné d'un démonstratif (29) ou d'un article défini (30), comme à Parakalamos :

(29) Romani (Matras 2004 : 93)

Ther-as-as	*but*	*buti*	*akala*	*duj*	*berš.*
posséder-1Pl.IMPF	beaucoup	travail	DEM.PL	deux	année(M.SG)

« On avait beaucoup de travail les deux dernières années. »

(30) Romani (Matras 2004 : 75)

O	*kher*	*e*	*duj-en-gor-o*	
ART.DEF.M.SG	maison(M.SG)	ART.DEF.PL.OBL	deux-PL-GEN-M.SG	
e	*phral-en-gor-o*	*isi*	*but*	*čik-or-o.*
ART.DEF.PL.OBL	frère-M.PL-GEN-M.SG	être.3.PRES	très	petit-DIM-M.SG

« La maison des deux frères est très petite. »

L'exemple (30) présente une étonnante structure DEF N [DEF NUM DEF N]^{GEN}. Le syntagme génitival DEF NUM DEF N donne l'impression d'une réduplication de l'article défini – ce qui laisse penser que le numéral a ici la fonction d'un adjectif. DEF NUM DEF N serait l'équivalent de DEF A DEF N, syntagme polydéfini avec adjectif antéposé – qui est par ailleurs possible dans cette variété (exemple (61). D'autres exemples le confirment :

(31) Romani (Matras 2004 : 75)

Kam-am-as	*te*	*dž-av-as*	*ti*	*poli*
vouloir-1SG-IMPF	COMPL	aller-1SG-IMPF	à.ART.DEF.F.SG	ville(F.SG)
ol-en-za	*i*	*trin-en-za*	*i*	*murš-en-za*
DEM-PL.OBL-INSTR	ART.DEF.PL.OBL	trois-PL.OBL-INSTR	ART.DEF.PL.OBL	homme-M.PL.OBL-INSTR

« Je voulais aller en ville avec ces trois hommes. »

(32) Romani (Matras 2004 : 75)

Šun-d-om	*ol-es-ke*	*m-i*	*duj-en-dar*	*mi*
entendre-PRET-1SG	DEM-M.SG.OBL-DAT	POSS.1SG-PL.OBL	deux-PL.OBL-ABL	POSS.1SG-PL.OBL
phral-en-dar.				
frère-PL.OBL-ABL				

« J'ai entendu parler de lui par mes deux frères. »

L'exemple (31) est similaire à l'exemple (30) : le syntagme DEM DEF NUM DEF N présente deux articles définis, l'un avant le numéral, l'autre avant le nom – à ceci près que la présence d'un démonstratif alourdit encore la détermination. L'exemple (32) confirme l'hypothèse de la réduplication : pas d'article défini dans ce syntagme POSS NUM POSS N, mais un possessif antéposé au numéral et réduplicqué. On a vu en section 1.3.4.2, p. 126, que cette variété postposait le possessif lorsque le nom est accompagné d'un adjectif ou d'un numéral (Matras 2004 : 76) : *me duj me phenja*, mes deux sœurs (POSS NUM POSS N), *mo baro mo phral*, mon grand frère (POSS A POSS N). Notons que l'arli d'Albanie présente la même structure DEF NUM DEF N que la variété de Parakalamos :

(33) *Romani (F65A, Korçë, 1er août 2014)*

O	*šov*	*o*	*saati*[1]	*čhid-asa*	*drom.*
ART.DEF.PL	six	ART.DEF.PL	heure(M.SG)	partir-1SG.PRES	chemin

« A six heures on est partis. »

1.3.6. Syntagme génitival

1.3.6.1. Forme

Le syntagme génitival est constitué d'un nom fléchi au génitif, portant la marque -*k*(*er*)- suivie d'une désinence adjectivale qui s'accorde en genre et en nombre avec le nom-tête. Dans de nombreuses variétés, le morphème -*ker*- présente une variation vocalique (-*kir*-, -*kor*-, -*ker*-) que Matras (2004 : 73) qualifie d'introflexion : la désinence adjectivale, -*i*, -*o*, -*e*, est copiée dans le segment vocalique interne du morphème du génitif -*kVr*-. Friedman (1993 : note 34) esquisse une description de la variation morphologique du génitif et de son usage dans quelques variétés de Macédoine. Le nom fléchi au génitif est en général accompagné d'un déterminant (l'article défini *e* dans l'exemple (38) voire de modifieurs propres. C'est cet ensemble (éventuel déterminant au cas oblique + éventuel adjectif au cas oblique + nom fléchi au génitif + éventuel autre modifieur postnominal) que j'appellerai « syntagme génitival ».

Il existe une différence d'emploi entre génitif long (morphème -*k*V*r*-, où la voyelle est assimilée à celle qui suit, c'est-à-dire à la désinence adjectivale, ici /o/), et court (morphème -*k*- seul), illustrée par l'exemple (36). Seul le génitif « long » peut revêtir la fonction syntaxique d'attribut, de l'objet en (36) et de sujet en (27). L'emploi épithète (ancrant ou non, antéposé ou postposé) peut au contraire se faire tant avec la forme courte (exemples (35) ou (36)) qu'avec la forme longue (exemple (37)). (Matras 2002 : 89) rappelle que Sampson avait noté cette scission fonctionnelle il y a presque un siècle (Sampson 1926 : 86-88).

Le syntagme génitival du romani constitue un défi typologique (Koptjevskaja-Tamm 2000 : 147). Il s'agit d'un syntagme nominal ayant pour tête un nom, éventuellement accompagné de son déterminant et de ses modifieurs. Ce nom étant fléchi au cas fusionnel oblique (strate 1), l'ensemble du syntagme est régi par ce même cas. Le nom est le seul à porter un morphème agglutinant, qui entre en relation paradigmatique avec les autres cas de la strate 2, à l'origine des postpositions clitiques. Il porte enfin un morphème fusionnel de type adjectival (strate 1), qui le fait s'accorder morphologiquement à son nom-tête, exactement comme un adjectif. Ces deux derniers morphèmes font du génitif romani, comme de celui de la plupart des autres langues indiennes modernes, un cas de *Suffixaufnahme*. Plank (1995 : V) définit ce phénomène comme « *the use of case (plus perhaps further inflectional categories, in particular number) for purposes of agreement, linking a noun*

[1] L'expression de l'heure est soumise à variation : le nom pourrait ici être mis au pluriel (*saatja*), de même que le syntagme nominal pourrait être introduit par une préposition (*t*-) : *to deš to saatja*, à dix heures (PREP.DEF NUM PREP.DEF N) (syntagme prononcé par le frère de cette locutrice).

already carrying a case (plus perhaps number), typically a genitive, to another nominal to which it is syntagmatically related, typically (or so it often seems) as an attribute ».

Le syntagme genitival est, en romani, antéposé à la tête, de même que dans les autres langues indiennes (Friedman 1993 : 20-23). Il explique cependant que les variétés sud-balkaniques présentent une grande variation sur ce point, en raison du contact avec les différentes langues locales : l'albanais postpose toujours le syntagme génitival ; le macédonien et le bulgare le postposent, mais acceptent l'antéposition (considérée comme un turcisme) ; le turc l'antépose. Les variétés de romani en contact avec le slave acceptent la postposition du syntagme génitival par rapport à la tête. En Macédoine, la postposition est, à l'époque où écrit l'auteur, caractéristique du style littéraire (Friedman 1993 : 23).

Il est le plus souvent considéré comme un modifieur, au même titre qu'un adjectif ou qu'une proposition relative. D'un point de vue sémantique et fonctionnel, le syntagme génitival relève de deux types sémantiques : ancrant/référentiel ou bien non-ancrant/non-référentiel/relationnel. Le génitif non-ancrant peut être un procédé très productif de nouveaux vocables (adjectifs, voire substantifs). Certaines variétés assignent une position, antéposée ou postposée au nom-tête, à chacun des deux types – mais pas toutes. Lyons (1985, 1986) distinguent deux types de génitifs dans les langues du monde :

- *adjectival-genitive* : le syntagme génitival est placé dans une position réservée à l'adjectif (Lyons 1999 : 24). Il peut, tel un adjectif, avoir une fonction d'attribut du sujet, d'attribut de l'objet ou de modifieur épithète, antéposé ou postposé au nom-tête. On l'appelle également génitif adnominal : un syntagme fonctionnant comme constituant d'un syntagme nominal et dont la relation au nom-tête est marquée par le génitif.
- *determiner-genitive* : le syntagme génitival est placé dans une position réservée au déterminant défini et jouant ce rôle (Lyons 1999 : 24). Il peut revêtir la fonction de déterminant, à la condition d'être antéposé au nom-tête et de type ancrant. Dans les langues germaniques, on l'appelle également génitif saxon : une expansion au génitif d'un nom-tête, antéposé à celui-ci et qui est incompatible avec la présence d'un déterminant. Une telle structure implique toujours que le nom-tête ait une valeur définie.

Nous allons voir, dans les deux sections suivantes, que le romani est doté de ces deux types.

1.3.6.2. Fonction de modifieur adnominal

Type 1 « ancrant » / « référentiel »

Koptjevskaja-Tamm (2000 : 125-128), suivant les travaux de Hawkins et de Fraurud, définit la fonction ancrante comme l'identification du référent du nom-tête. C'est la fonction principale de la plupart des syntagmes génitivaux : servir de point de référence pour le nom-tête. Si l'on sait qui est *Maxime*, on peut identifier « la chaussure de Maxime » ou « l'idée de Maxime ». Koptjevskaja-Tamm (2000 : 126) dresse une liste de relations possibles entre les deux noms : liens de parenté, parties du

corps, relation à une personne ou à un animal, relations métonymiques, relation de possession, états, procès… Dans l'exemple (34), le nom au génitif, *Demir Kosturi*, est une entité du monde qui permet de distinguer plus précisément la *famille* dont il est question. Le référent du nom-tête *famille* est identifié via sa relation au référent du nom au génitif, *Demir Kosturi*.

(34) *Romani (M52A, Korçë, 25 juillet 2014)*

Lesk-o	*nip*	*isi…*	*o*	Arben	Kosturi,
POSS.3SG.M-M.SG	petit-fils(M.SG)	être.3.PRES	ART.DEF.M.SG	Arben	Kosturi,
o	*Kostaqi*	*em*	*o*	*Yli.*	
ART.DEF.M.SG	Kostač	et	ART.DEF.M.SG	Ylli	
ti	*familja*	*te*	*Demir*	*Kostur-es-ɛ-i.*	
dans.ART.DEF.F.SG	famille(F.SG)	dans.ART.DEF.M.SG.OBL	Demir	Kosturi-M.SG.OBL-GEN-F.SG	

« Son petit-fils est… Arben Kosturi, Kostač et Yli. Dans la famille de Demir Kosturi »

Type 2 « non-ancrant » / « non-référentiel »

Koptjevskaja-Tamm (2000 : 140-144) définit la fonction non-ancrante comme la qualification ou la classification du nom-tête. Le syntagme génitival exprime une propriété, de la même manière qu'un adjectif : âge, taille, matériau, mesure, similarité, caractéristique temporelle ou spatiale… Koptjevskaja-Tamm (2000 : 143). Il ne fait donc pas référence à une entité spécifique. Le syntagme génitival est la manière la plus courante d'indiquer l'âge en romani :

(35) *Romani (F65A, Korçë, 1er août 2014)*

Šaj	*te*	*vak'j-av*	*t'-ke,*	*bišupandž-e*
possible	COMPL	dire-1SG	2SG.OBL-DAT	vingt_cinq-M.PL.OBL
breš-en-g-o	*raklo.*	*But*	*tern-o.*	
année-M.PL.OBL-GEN-M.SG	garçon_non_rom(M.SG)	beaucoup	jeune-M.SG	

« Je peux te dire, un garçon de vingt-cinq ans. Très jeune. »

En (35), le nom au génitif (« année ») n'est pas spécifique, contrairement au nom-tête (« garçon »). Ce n'est même pas le cas en (36) pour « plat » et « viande ». Dans cet exemple, « de chou » et « de bœuf » ne font référence à aucun chou ou bœuf spécifiques. « Plat au chou » et « viande de bœuf » constituent des types de plat et de viande, où « de chou » et « de bœuf » ont le sens générique d'adjectifs tels que « chou-esque » et « bovin ».

(36) *Romani (F85A, Korçë, 23 juillet 2014)*

Xabe	*ker-asa*	*amen*	*šah-es-kor-o.*
plat(M.SG)	faire-1PL	1PL	chou-M.SG.OBL-GEN-M.SG
Mas-e-sa.	*Kaur-us-k-o*	*mas.*	
viande-M.SG.OBL-INSTR	boeuf-M.SG.OBL-GEN-M.SG	viande(M.SG)	

« Comme plat, nous on en fait au chou. Avec de la viande. De la viande de bœuf. »

Le syntagme génitival joue le rôle d'un adjectif dit relationnel. En français également, les adjectifs relationnels sont dérivés de noms : *universitaire* de *université*, *économique* de *économie*, et *bovin* de

bœuf (Dubois et al. 2012 : 17). Un tel syntagme génitival est alors en concurrence avec ces adjectifs relationnels dérivés en *-(v)alo, -ano, -ikano, -no, -u(t)no, -av/mno* : Matras (2002 : 77) cite l'opposition relevée pour « plume d'oiseau » par Halwachs (1998 : 107), entre *čiriklano por*, plume 'oiseline', et *čiriklakero por*, plume d'oiseau. L'ensemble est équivalent d'un nom composé dans les langues où ce procédé est productif (comme l'allemand). Le romani, lui, ne présente quasiment aucun mot composé formé par juxtaposition de deux noms, à quelques exceptions dialectales que relève (Matras 2002 : 76) : *phrala phena*, frères et sœurs (litt. frères sœurs), *kańhajaro*, œuf (litt. poule œuf) à Šóka et Farkašda, *šéro róm*, chef rom (litt. tête rom) en Pologne.

Distribution des deux types

Comme l'adjectif, la position canonique (et héritée) du syntagme génitival est antéposée. Matras (2002 : 76) explique que l'ordre normal des constituants est celui où le syntagme génitival précède son nom-tête, de même que l'adjectif précède normalement le nom. L'antéposition du syntagme génitival constitue un héritage de la syntaxe nominale indo-aryenne, alors que le syntagme verbal a changé au profit d'un ordre majoritairement VO. Cette tension rend le romani typologiquement hybride selon Matras (2002 : 166), et explique que certaines variétés poursuivent le changement syntaxique, c'est-à-dire assouplissent l'ordre des constituants nominaux. Le syntagme génitival peut alors être postposé, occupant l'emplacement postnominal que l'auteur qualifie d'« optionnel ». Comme pour l'adjectif, donc, une certaine souplesse est possible, puisqu'il peut être postposé au nom dans certains contextes et dans certaines variétés. Selon Boretzky (1994a : 117), le placement est libre en kalderaš, mais la postposition légèrement plus fréquente que l'antéposition. On préfère toutefois antéposer le syntagme génitival s'il est accompagné de modifieurs ; on évite au contraire de l'antéposer lorsque le nom-tête présente, lui, d'autres modifieurs : « l'anniversaire de la fille non-rom » se dit indifféremment *o rodjendano* <u>*e rakhljako*</u>, litt. l'anniversaire de la fille non-rom, ou <u>*e rakhljako*</u> *rodjendano*, litt. de la fille non-rom anniversaire, mais « le vingtième anniversaire de la fille non-rom » se dit seulement *o bišto rodjendano* <u>*e rakhljako*</u>. C'est le cas dans la variété de Parakalamos : *o kher* <u>*e dujengoro e phralengoro*</u> *isi but čikoro*, la maison des deux frères est très petite (Matras 2004 : 75). Dans d'autres au contraire, le syntagme génitival occupe une place fixe : Tenser (2005 : 54-55) et Cech (2006 : 43) notent pour les variétés de Lituanie et de Slovénie qu'il est antéposé, la postposition pouvant s'interpréter comme un « procédé poétique » de déplacement. En revanche, le syntagme génitival est postposé dans l'écrasante majorité des cas à Parakalamos : *o kher* <u>*e muršeskoro*</u> *isine but purano*, la maison <u>de l'homme</u> était très ancienne (Matras 2004 : 97).

La position du syntagme génitival peut avoir une fonction sémantique, comme en lovari où s'opposent *kher* <u>*le dilengo*</u>, asile (litt. maison <u>des fous</u>), dans le sens générique de « maison de fous », et <u>*le dilengo*</u> *kher*, maison des fous (litt. <u>des fous</u> maison), dans le sens spécifique de « maison appartenant aux personnes qui sont folles » (Matras 2002 : 77, 166). Le premier syntagme génitival, postposé, a une fonction non-ancrante, générique, tandis que le second, antéposé, a une fonction ancrante, spécifique.

Les exemples de notre corpus albanais ne correspondent toutefois pas à ce modèle : les exemples (35) et (36), non-ancrants, sont antéposés, tandis que l'exemple (34), ancrant, est postposé. On trouve d'ailleurs des syntagmes génitivaux non-ancrants qui sont postposés :

(37) *Romani (F60A, Korçë, 21 juillet 2014)*

Vakjer-ena	*les-ɕe*	*čomleči.*	*Čomleči*	*purum-a-kor-o.*
dire-1PL	3SG.M.OBL-DAT	*çomlek*	*çomlek*	oignon-F.SG.OBL-GEN-M.SG

« On l'appelle çomlek. Çomlek à l'oignon. »

Le syntagme génitival en (37), même postposé, garde ici sa fonction de modifieur épithète et n'est pas à proprement parler optionnel. Il a le statut d'un adjectif relationnel et construit, avec le nom-tête, un type de plat. Cela amène selon nous à remettre en question l'ordre considéré comme canonique des constituants nominaux.

1.3.6.3. Fonction de déterminant

Les arguments en faveur de la fonction de déterminant sont d'ordre distributionnel. Dans l'exemple (38), le syntagme « de l'organisation » est placé après le nom-tête « Rom ». On a donc une structure DEF A N [DEF N]^{GEN} avec deux déterminants.

(38) *Romani (M46A, Korçë, 25 juillet 2014)*

Tu	*isin-an*	*o*	*šerutn-o*	*rom,*
2SG	être-2SG	ART.DEF.M.SG	principal-M.SG	rom
o	*bar-o*	*rom…*	*e*	*organizat-a-k-o.*
ART.DEF.M.SG	grand-M.SG	rom	ART.DEF.F.SG.OBL	organisation-F.SG.OBL-GEN-M.SG

« Toi tu es le chef, le grand Rom… de l'organisation. »

Lorsque le syntagme génitival est antéposé, le déterminant du nom-tête disparaît. Au lieu d'une structure POSS [DEF N]^{GEN} N, on obtient [POSS N]^{GEN} N comme en (39), au lieu de DEF [DEF N] ^{GEN} N, on obtient [DEF N]^{GEN} N comme en (40) :

(39) *Romani (F16AČ, Korçë, 25 juillet 2014)*

Akova	*isi*	*me…*	*me*	*daj-a-k-o…*	*phral,*	*me*	*daj-a-k-o…*	*phral.*
DEM	être.3.PRES	POSS.1SG.OBL	POSS.1SG.OBL	mère-F.SG.OBL-GEN-M.SG	frère	POSS.1SG.OBL	mère-F.SG.OBL-GEN-M.SG	frère

« Lui c'est… le frère de ma… de ma mère…, le frère de ma mère. »

(40) *Romani (F53A, Korçë, 25 juillet 2017)*

Amar-i	*čhib?*	*ẽ?*	*apo*	*e*	*Ben-es-ɕ-i*	*čhib?*
POSS.1PL-F.SG	langue(F.SG)	hein	ou	ART.DEF.M.SG.OBL	Beni-M.SG.OBL-GEN-F.SG	langue(F.SG)

« Notre langue ? hein, ou la langue de Beni ? »

L'article *e* dans cette expression est au cas oblique, ce qui montre qu'il détermine le nom *Beni* au génitif, et non *čhib*, langue, au cas direct. Ce dernier nom n'a donc pas de déterminant propre. Ainsi la présence d'un syntagme génitival est-elle incompatible avec la présence d'un déterminant défini

portant sur le nom-tête. Koptjevskaja-Tamm (2000 : 128-129) montre que cette incompatibilité existe également dans d'autres langues, comme l'anglais, le gaélique d'Ecosse, le maltais ou le suédois. Si l'on postule que le syntagme génitival occupe la place syntaxique d'un adjectif, car il a la fonction de modifieur, alors la place du déterminant est laissée vide. Or on a vu, pour le démonstratif (1.3.3.2) et le possessif (1.3.4.2), que le romani ne supporte pas un tel vide : c'est pourquoi on peut supposer que c'est le syntagme génitival lui-même qui vient le combler. Une autre raison est sémantique : s'il y avait véritablement article zéro dans l'exemple (40), alors le nom « langue » serait indéfini (section 1.3.2.2, p. 119). Or il a bel et bien un sens défini grâce au syntagme génitival. Lorsque celui-ci est de type ancrant, il permet d'identifier le référent en apportant des éléments d'information sur le nom-tête. On peut donc dire que le trait de définitude se situe sur le syntagme génitival. Lorsque celui-ci est de type non-ancrant, il apporte une information sur le nom-tête, sans pour autant le rendre défini. Il reste alors modifieur et peut cohabiter avec *jekh*[1] :

(41) Romani (Koptjevskaja-Tamm 2000 : 132)[2]

jekh-e	*Petr-os-k-ə*	*čav-ex-ke* [sic]
un-M.SG.OBL	Pierre-M.SG.OBL-GEN-M.SG.OBL	Pierre-M.SG.OBL-DAT

« pour un enfant de Pierre »

Il existe cependant des langues où le génitif n'est pas incompatible avec l'article du nom-tête : c'est le cas du romani lorsque le syntagme génitival est postposé ou qu'il est antéposé de type non-ancrant, mais aussi du bulgare ou du catalan (Koptjevskaja-Tamm 2000 : 129). On a vu (section 1.3.6.1, p. 130) que les langues du monde présentent deux types de génitifs : *determiner-genitive* et *adjectival-genitive* :

- Lorsque le syntagme génitival est postposé au nom, il occupe la place et la fonction d'un adjectif et relève du type *adjectival-genitive* : *i zor le Devleski*, le pouvoir de Dieu (Gjerde et Kristiansen 1994 : 130).
- Lorsque le syntagme génitival est antéposé au nom et qu'il est de type ancrant, il occupe la place et la fonction d'un déterminant défini et relève du type *determiner-genitive* : *le Devleski zor*, le pouvoir de Dieu (Gjerde et Kristiansen 1994 : 128).

Cette particularité semble héritée, du moins n'est-elle pas empruntée au grec. Ce dernier postpose le plus souvent le syntagme génitival (DEF N [DEF N]^{GEN}), mais lorsqu'il l'antépose, il conserve l'article défini du nom-tête ([DEF N]^{GEN} DEF N) :

(42) Grec (Boretzky 1994a : 165)

τῆς	μάνα-ς	τό	σπίτι
ART.DEF.F.SG.GEN	mère-F.SG.GEN	ART.DEF.N.SG	maison[N.SG]
τό	σπίτι	τῆς	μάνα-ς

[1] Koptjevskaja-Tamm (2000 : 132) remarque que *jekh*, dans de tels cas, fonctionne comme un article ou comme un numéral. Dans ces deux cas, il occupe la place et la fonction d'un déterminant, au détriment du syntagme génitival.

[2] Données issues d'un questionnaire typologique en mečkar, du čergar (d'Albanie ?), du kalderaš et de l'erli de Bulgarie (Koptjevskaja-Tamm 2000 : 147)

ART.DEF.N.SG	maison[N.SG]	ART.DEF.F.SG.GEN	mère-F.SG.GEN

« la maison de la mère »

On constate que le grec tolère que le syntagme nominal commence par un modifieur du nom, et que le déterminant puisse donc suivre le modifieur – ce qui n'est pas le cas du romani (section 1.1, p. 103). Boretzky (2000b : 42) note cependant quelques exceptions, où le syntagme génitival est antéposé à la fois au nom-tête et à son déterminant ([DEF N]^{GEN} DEF N) :

(43) Grec (Boretzky 2000b : 42)

τῶν	ἀνθρώπ-ων	τὰ	μάτι-α
ART.DEF.M.PL.GEN	homme-M.PL.GEN	ART.DEF.N.PL	œil-N.PL

Romani[1] (Boretzky 2000b : 42)

i [sic]	*džen-en-ger-e*	*o*	*jakh-a*
ART.DEF.M.PL.OBL [?]	personne-M.PL.OBL-GEN-PL	ART.DEF.PL	œil-MPL

« les yeux des gens »

Selon Boretzky (2000b : 42), cette construction se trouve « dans de nombreux dialectes », y compris des dialectes sporadiques « pour lesquels on peut exclure une forte influence du grec ». Il est donc possible qu'elle soit héritée – elle correspond d'ailleurs au schéma de l'ordre des constituants dans les langues indo-aryennes modernes, proposé par Masica (1991 : 373) (section 1.1, p. 103). Nous ne l'avons cependant pas entendue en Albanie.

1.4. L'adjectif

1.4.1. Flexion

Les adjectifs connaissent les mêmes catégories que le nom et s'accordent avec lui en genre, nombre, cas (strate 1) et animéité. Notons que les adjectifs ne s'accordent pas en définitude, comme ce peut être le cas dans les langues sémitiques centrales (Rubin 2005a : chap. 4, Rubin 2005a : 62), par exemple en hébreu (Alexiadou 2005 : section 3.4). Il existe deux classes morphologiques d'adjectifs hérités :

- une classe majoritaire distinguant le genre au singulier[2] (*bari*, grande, *baro*, grand)
- une classe minoritaire ne le distinguant pas[2] en raison de la finale consonantique (*godjaver*, intelligent)

Des adjectifs empruntés de longue date, notamment au grec, suivent un paradigme particulier. Ils ne distinguent pas le genre au singulier ni au pluriel, comme les adjectifs hérités à finale consonantique, mais présentent un affixe emprunté au grec, *-on-*, entre la racine et la désinence de cas oblique (Matras

[1] L'auteur ne précise d'ailleurs pas de quelle variété est issu son exemple.

[2] Le genre est indistinct au pluriel (cf. section 1.4.1, p. 38).

2002 : 95). C'est le cas d'adjectifs courants tels que *lungo*, long (roumain), *drago*, cher (slave du sud), mais aussi de ceux portant le suffixe *-it(i)ko* / *-icko*, tels que *krajasicko*, royal (Boretzky 1994a : 48). D'autres adjectifs, empruntés récemment aux langues de contact, continuent à suivre le paradigme de la langue en question, ainsi *dosadni*, ennuyeux.PL, *jednaki*, égaux, accordés au masculin pluriel selon le paradigme serbo-croate (Boretzky 1994a : 48). L'accent tonique des adjectifs empruntés est conservé sur la syllabe de la langue d'origine : *lúngo*, *drágo*.

Dans certains dialectes, l'adjectif est fléchi à la strate 1 tout comme à la strate 2, à l'instar d'un nom. C'est le cas du romani de Parakalamos (Matras 2004 : 74-75) : « dživdilom jek bareste khereste *'I lived in a big.LOC house.LOC'*, tho te vast tatesa panjesa *'wash your hands with hot.INSTR water.INSTR'*. » C'est également le cas lorsque l'adjectif modifie un nom au génitif (exemple (30). On note que l'adjectif est antéposé dans ces structures DEF A DEF N, IND A N et Ø N. Cech et Heinschink (2001 : 159) signalent que les adjectifs et les numéraux qui précèdent un nom sont eux-mêmes fléchis comme des noms : « poraneste dadeste '*at grandfather's' (locative),* barvalaha romnjaha *'with a rich woman' (instr.)* ». Il s'agit de structures Ø N.

1.4.2. Degrés

1.4.2.1. Comparatif

Certaines variétés ont conservé l'ancien comparatif synthétique formé à l'aide du suffixe -eder, comme le romani de Lituanie (Tenser 2005 : 13) ou du Burgenland autrichien (Halwachs et al. 2002 : 23). On le rencontre dans certaines variétés des Balkans pour des adjectifs hérités très courants tels que *terno*, jeune, ou *lačho*, bon, (kalderaš du Banat, Boretzky 1993 : 42). Le degré comparatif des adjectifs et des adverbes est pour le reste construit à l'aide d'une particule invariable antéposée à l'adjectif[1] – il en existe plusieurs dans les Balkans :

- *mo*, de l'albanais *më*, dans la variété grecque de Parakalamos (Matras 2004 : 75),
- *daa*, du turc *daha*, ou *džin*, hérité, dérivé de l'adverbe (*a*)*dži* / *adžaj*, dans la variété grecque d'Ajios Athanasios (Sechidou 2011 : 34),
- *daha*, du turc *daha*, dans la variété turque des Sepečides (Cech et Heinschink 1999 : 26),
- *pio*, du grec πιό, pió, dans la variété grecque d'Epire (Matras 2002 : 203),
- *po*, des macédonien et bulgare по-, *po*, dans la variété burgudži du Kosovo (Boretzky 1993 : 42),
- *maj*, du roumain en kalderaš du Banat (Boretzky 1994a : 48), dans les dialectes centraux (Matras 2002 : 78), ainsi que dans toutes les variétés *vlax*, comme le romani de Bogota (Deman 2005 : 79) ou du Mexique (Adamou 2013 : 1090-1091, 1094),
- *naj*, du slave dans certaines variétés des Balkans, du centre et du nord-est (Matras 2002 : 78).

[1] Il s'agit d'un trait commun aux variétés balkaniques du romani, et aux langues balkaniques en général.

Le čergar I albanais utilise *maj*, emprunté au roumain. L'arli et le mečkar utilisent *po*, qui peut être emprunté au macédonien, bulgare, ou bien au grec πιό, pió :

(44) *Romani (M46A, Tiranë, 8 août 2014)*

Tu	*te*	*inkljo-š*	*po*	*šukar*	*kato*	*vaver-a.*
2SG	COMPL	sortir-2SG	COMP	bien	venant_de.ART.DEF.PL	autre-PL

« Pour que tu t'en sortes mieux que les autres. »

1.4.2.2. Superlatif

Le degré superlatif des adjectifs et des adverbes se forme en faisant précéder le comparatif (adjectif-*eder* ou particule + adjectif) de l'article défini : voir Matras (2004 : 75) pour la variété de Parakalamos, Sechidou (2011 : 34) pour la variété d'Ajios Athanasios, Boretzky (1994a : 48) pour le kalderaš du Banat). C'est le cas de l'arli et du mečkar, comme dans cet exemple mečkar où une locutrice liste ses frères et sœurs :

(45) *Romani (F47M, Llakatund, 13 août 2014)*

Isin-om	*ə*	*Drita.*	*I*	*Valide*	*ə*	*po*	*tikin-i.*
être-1SG.PRES	ART.DEF.F.SG	Drita(F.SG)	ART.DEF.F.SG	Valide(F.SG)	ART.DEF.F.SG	COMP	petit-F.SG

« Moi je suis Drita. Valide [est] la plus petite. »

Dans les variétés où l'article défini est obligatoire, le superlatif ne se distingue donc pas du comparatif : *i po tikini*, qui signifie « la plus petite », est ambigu en romani comme en français.

Les variétés du nord-est, du centre-nord et certaines des Balkans n'emploient pas l'article défini mais une particule de superlatif, *naj-* empruntée au slave *naj-* (Matras 2002 : 203). En burgudži (Boretzky 1993 : 42), on emploie l'article défini + *naj-* + adjectif (au comparatif ou non). La variété du Burgenland autrichien a le choix entre la particule *lek*, empruntée au hongrois, ou bien les particules (plus rares) *naj-* et *maj-*, empruntées au slave, placées avant l'adjectif dérivé au comparatif (Halwachs et al. 2002 : 23).

1.4.3. Distribution

La position de l'adjectif par rapport au nom sur l'axe syntagmatique exprime son rapport au nom-tête du syntagme nominal.

1.4.3.1. Antéposition canonique

L'adjectif est canoniquement anteposé dans toutes les langues indo-aryennes modernes selon Masica (1991 : 370) : « *Immediately to the left of the Head Noun come attributive Adjectives, if any: B.*

dhārmik mohilā *'religious woman'*, *N.* purāno mandir *'old temple'*, *Si.* sudu eladenā *'(the) white cow'* ». Matras (2002 : tableau 7.1 p166) le confirme pour le romani, de même que la description de nombreuses variétés individuelles, par exemple :

- le romani de Lituanie (Tenser 2005 : 54),
- le romani de Thrace collecté et étudié par Adamou (2016 : 161) : 38 occurrences d'adjectifs antéposés, quel que soit le déterminant introduisant le syntagme. La postposition de l'adjectif y est très minoritaire avec seulement quatre occurrences, que l'autrice qualifie de marquées.
- le romani d'Ajia Varvara (Igla 1996 : 164-166),
- le romani de Parakalamos : L'adjectif des syntagmes introduits par un article indéfini y est presque toujours antéposé (*nje baro džukel*, un gros chien) (Matras 2004 : 97).

Dans d'autres variétés, la position canonique est postposée au nom, ainsi :

- le romani des Abruzzes (Soravia 1972 : 38),
- le romani *vlax* de Bogota, où la position non-marquée de l'adjectif est postposée dans une structure polydéfinie (DEF N DEF A). La position antéposée (DEF/IND A N) est marquée (cf. section 2.3, p. 447).

1.4.3.2. Pourquoi postposer l'adjectif ?

Comment expliquer ces divergences ? La syntaxe de l'adjectif est beaucoup moins étudiée que sa morphologie (flexion, dérivation…). Quelques travaux notent cependant la possibilité de postposer l'adjectif :

1) **Caractère de commentaire** : c'est en quelque sorte la fonction de l'adjectif épithète lorsque le syntagme nominal est introduit par un article indéfini selon Boretzky (1993 : 41) :

> « *Das attributive Adj. kann auch auf das Nomen folgen. Es findet sich eher in inderterminierten NPs und hat dann Kommentcharakter ; vgl.* jek phuro koro *„ein blinder Alter“,* manuš durutno sjum *„ich komme von weit her“,* jek phabaj lačhi *„ein schöner Apfel“. In dieser Position flektiert das Adj. normalerweise voll; vgl.* vakerla pe phraleske cikoreske *„er sagt zu seinem jüngeren Bruder“;* me phraleske (e) baredereske *„meinem größeren Bruder“;* me džamutres najphurederes *„meinen ältesten Schwiegersohn“* » (Boretzky 1993 : 41)

L'auteur ajoute trois exemples qui sont cependant introduits par un possessif, partant définis. La flexion nominale (strates 1 et 2) est alors « copiée » du nom à l'adjectif :

Romani (Boretzky 1993 : 41)

(46)	*vaker-l-a*	*p-e*	*phral-es-ke*	*cik-or-es-ke*
	parler-PRET-3SG	POSS.REFL-M.SG.OBL	frère-M.SG.OBL-DAT	petit-DIM-M.SG.OBL-DAT

« Il dit à son jeune frère. »

(47)	*m-e*	*phral-es-ke*	*(e)*	*bar-eder-es-ke*

POSS.1SG-M.SG.OBL frère-M.SG.OBL-DAT ART.DEF.M.SG.OBL grand-COMP-M.SG.OBL-DAT

« à mon frère aîné »

(48) *m-e* *džamutr-es* *naj-phur-eder-es*

POSS.1SG-M.SG.OBL gendre-M.SG.OBL SUP-vieux-COMP-M.SG.OBL

« mon gendre le plus âgé »

En principe, l'adjectif ne se fléchit qu'à la strate 1 (section 1.4.1, p. 136) : sa forme oblique est *-e*. Or il porte en (48) la désinence nominale *-es* (strate 1), en (46) et (47) les désinences *-es-ke* (strates 1 et 2). Cela rappelle le romani de Parakalamos (Matras 2004 : 74-75) où l'adjectif est fléchi comme un nom en lui étant toutefois antéposé. En (47), un article défini peut précéder l'adjectif postposé.

2) **Syntaxe souple** : Igla (1996 : 164-166) parle d'une très grande liberté des constituants au sein du syntagme nominal et d'une influence grecque dans la variété d'Ajia Varvara. Si l'adjectif est postposé dans un syntagme défini, il est obligatoirement accompagné d'un article défini (cf. section 2.1, p. 142). L'autrice n'explique pas dans quel contexte l'adjectif se postpose, mais précise qu'il suit dans ce cas la flexion nominale et non la flexion adjectivale. S'il y a deux adjectifs, ils encadrent le nom.

3) **Emplacement optionnel** : Postposé, l'adjectif appartient alors au *postnominal 'option' slot* selon Matras (2002 : 166-167) qui analyse deux exemples (Matras 2002 : 166-167) :

« *Discourse data provide some insights into the communicative triggers behind the placement of attributive elements in the postnominal 'option' slot:*

(3) muj i phuri, ačhile kadla rakle čore
died.F ART old.FEM remained.PL these boys poor.PL
(Bugurdži; Boretzky 1993: 87)
'the old woman died, these poor boys remained'

(4) ande jekh kesave cikno kheroro, žanes, kheroro
in indef such little house know.2SG house
cikno polski, tu žanes sar (Lovari)
small Polish you know.2SG how
'in such a small house, you know, a small Polish house, you know what kind'

In (3) the adjective čore *'poor' is exposed as an afterthought, evaluating the state of affairs referred to in the actual statement, i.e. the fact that the boys remained after the death of the old woman makes them qualify as 'poor'. In (4), the noun and its adjective* cikno *are repeated in order to add a further characterisation,* polski *'Polish', which here too can be argued to be evaluative, as it is expected to trigger an association on the part of the listener (since the listener is familiar with Polish houses, he will now understand the speaker's initial attempt to describe the house). The postnominal 'option' position is therefore a pragmatic position for most attributives, and a lexicalised position for genitives in some dialects. Individual dialects also show formal postnominal positions within the noun phrase, which are occupied either by calques (postposed demonstratives and adpositions) or by direct borrowings (postposed focus particles, such as Turkish-derived da).* » (Matras 2002 : 166-167)

Selon l'auteur, la structure ne peut s'expliquer qu'en discours, donc qu'en contexte : l'emplacement postnominal revêt une fonction spécifique, « pragmatique » et non lexicale, dans le cas des adjectifs transmettant un jugement, une évaluation, une caractérisation. L'aspect pragmatique provient de ce que le jugement est porté *a posteriori* : la fonction de l'emplacement postnominal est celle d'un *afterthought*. Cette notion, déjà employée par Matras (2000b : 102), n'est pas définie plus avant. Son exemple (4) est indéfini, comme les trois premiers exemples de Boretzky (1993 : 41) (cf. début de section) ou encore l'exemple d'Igla (1996 : 166) (*Inkaldás but šukár gilá nevé*, Elle a sorti de très belles nouvelles chansons). Deux adjectifs sont postposés : *cikno*, petit, employé plus haut dans la phrase, exprime une propriété déjà sous-entendue par le suffixe diminutif nominal *-oro*, et *polski*, polonais, qui exprime aussi une propriété de la maison, mais dont le locuteur semble vouloir faire un type de maison. Le premier adjectif étant donné actif, on peut penser que la postposition du deuxième sert à établir un lien avec le premier (« *to trigger an association on the part of the listener* » : sous-entendre peut-être que les maisons polonaises traditionnelles sont petites), à évoquer un savoir commun, la connivence avec l'interlocuteur. Un tel « commentaire après coup » joue dans ce cas un rôle extra- plutôt qu'intra-linguistique. C'est pourquoi l'auteur conclut que l'emplacement postnominal, qu'il qualifie d'option, est une position pragmatique. Les autres éléments que l'on y trouve sont des syntagmes génitivaux lexicalisés, des calques ou des emprunts non-intégrés.

2. La répétition de l'article défini

2.1. En romani

Le romani peut marquer plusieurs fois la définitude dans un même syntagme nominal :

(162) *Romani (F16AČ, Korçë, juillet 2014)*

Istanbul/	*kaj*	*ćer-ena*	*o*	*film-e*	*o*	*bar-e!*
Istamboul	où	faire-3PL.PRES	ART.DEF.PL	film-PL	ART.DEF.PL	grand-PL

« Istamboul, là où on fait les grands films ! »

La définitude est marquée deux fois, ici par deux articles définis qui accompagnent respectivement le nom « films » et l'adjectif « grands ». Par manque de place ou de temps, peu nombreuses sont les descriptions et études qui mentionnent ce phénomène, en raison de sa rareté. Il est donc possible de présenter ici celles qui le recensent, voire l'expliquent.

1) Boretzky (1994a : 55) remarque qu'un article défini peut être ajouté en **kalderaš du Banat serbe** lorsque le nom est déterminé par un démonstratif. Ce dernier peut être accompagné d'un article défini lorsqu'il est postposé (*e balen kodolen*, ces cochons, DEF N DEM) et lorsqu'il est antéposé (*kado kaš o baro*, ce grand arbre, DEM N DEF A). On constate que le nom est alors modifié par un adjectif :

(49) Romani (Boretzky 1994a : 55)

kado	*kaš*	*o*	*bar-o*
DEM.M.SG	arbre(M.SG)	ART.DEF.M.SG	grand-M.SG

« ce grand arbre, litt. cet arbre le grand »

2) Igla (1996 : 40, 85) mentionne elle aussi l'ajout de l'article défini en sus du démonstratif ou du possessif à **Ajia Varvara**. Elle décrit l'ordre des constituants nominaux (Igla 1996 : 166) :

> « *Adjektiva können vor dem head noun stehen oder ihm folgen; enthält die NP einen definiten Artikel, muß dieser bei Nachstellung des Adjektivs wiederholt werden – auch dies genau wie im Griechischen, vgl.:*
>
> | | | |
> |---|---|---|
> | *ART$_{indef}$ ADJ SUBST* | ekh barí čhej | μιά μεγάλη κοπέλα |
> | *ART$_{indef}$ SUBST ADJ* | ekh čhej barí | μιά κοπέλα μεγάλη |
> | | *„ein großes Mädchen“* | |
> | *ART$_{def}$ ADJ SUBST* | i barí čhej | ή μεγάλη κοπέλα |
> | *ART$_{def}$ SUBST ART ADJ* | i čhej i barí | ή μεγάλη ή κοπέλα |
> | | *„das große Mädchen“* » | |

Selon l'autrice, c'est la postposition de l'adjectif qui provoque la répétition de l'article défini, ce qui correspond à la structure du grec. On constate qu'elle ne mentionne pas l'existence, dans cette variété, de structure DEF N A (*i čhej barí*) – elle n'existe pas en grec. On constate également qu'elle met en

parallèle un syntagme DEF N DEF A romani (adjectif postposé) et un syntagme DEF A DEF N grec (adjectif antéposé)… alors que DEF N DEF A (ή κοπέλα ή μεγάλη) existe en grec. Elle ne mentionne pas, en revanche, l'existence en romani de *DEF A DEF N (**i barí i čhej*).

Igla (1996 : 166) note que l'adjectif postposé se décline soit comme un adjectif (50), soit comme un nom (51), c'est-à-dire selon la strate 2 :

(50) Romani (Igla 1996 : 166)

Dikh-l-ás	*e*	*rakl-és*	*e*	*bar-é*
voir-PRET-3SG	ART.DEF.M.SG.OBL	garçon_non_rom-M.SG.OBL	ART.DEF.M.SG.OBL	grand-M.SG.OBL

« Il vit le grand garçon non-rom. »

(51) Romani (Igla 1996 : 166)

Phen-d-ás	*e*	*bor-á-ke*	*e*	*cikn-á-ke*
dire-PRET-3SG	ART.DEF.F.SG.OBL	belle_fille-F.SG.OBL-DAT	ART.DEF.F.SG.OBL	petit-F.SG.OBL-DAT

« Il dit à la jeune belle-fille. »

3) Boretzky (2000b : 41), à la suite de Boretzky (1994a : 55), qualifie de *double* détermination (*double determination*)[1] les structures du type *e balen kodolen*, ces cochons, DEF N DEM. Boretzky (2000b : 42-43) qualifie de *sur*détermination (*overdetermination*)[2] les structures du type *kado kaš o baro*, ce grand arbre, DEM N DEF A (exemple (49). Ces dernières présentent deux articles définis et, d'une manière générale, deux déterminants portant sur un nom modifié par un adjectif qu'il dit « focalisé » (*focused*). Il oppose les exemples (52) et (53) en **bugurdži** :

(52) Romani (Boretzky 1993 : 171)

vaker-l-a	*p-e*	*phral-es-ke*	*cik-or-es-ke*
parler-PRET-3SG	POSS.REFL-M.SG.OBL	frère-M.SG.OBL-DAT	petit-DIM-M.SG.OBL-DAT

« Il dit à son petit frère »

(53) Romani (Boretzky 1993 : 171)

m-e	*phral-es-ke*	*e*	*cik-eder-es-ke*
POSS.1SG-M.SG.OBL	frère-M.SG.OBL-DAT	ART.DEF.F.SG.OBL	petit-COMP-M.SG.OBL-DAT

« à mon plus petit frère » (« *stress upon the adjective* » selon Boretzky 2000b:42)

Boretzky (2000b : 43) joint à ces exemples un cas de surdétermination d'une nature différente, la simple expression analytique du comparatif/superlatif (cf. sections 1.4.2.1, p. 137 et 1.4.2.2, p. 138) :

(54) Romani (Boretzky 1993 : 171)

o	*stəpîno*	*o*	*maj-bar-o*
ART.DEF.M.SG	maître(M.SG)	ART.DEF.M.SG	COMP/SUP-grand-M.SG

« le plus grand maître »

[1] Il s'agit d'un double marquage de la définitude via un *simple definite* et un *complex definite* (Lyons 1999).

[2] On peut ici penser à une influence de la linguistique scandinave, qui utilise ce terme pour qualifier le double marquage de la définitude (norvégien *overbestemthet).*

On constate que, contrairement à Igla (1996 : 166), il considère possible la structure DEF N A, puisque la répétition de l'article "peut" apparaître, mais ne le doit pas nécessairement. Le contexte déclenchant son apparition est la postposition de l'adjectif et le focus porté sur l'adjectif. Il est en outre le premier à noter que la surdétermination ne met pas nécessairement en jeu deux articles définis (DEM N DEF A, exemple (54), d'ailleurs contestable), mais deux déterminants (DEM N DEF A, exemple (49), ou POSS N DEF A, exemple (53)). Il n'y a donc pas à proprement parler répétition de l'article défini, comme l'écrivait Igla (1996 : 166), mais ajout d'un article défini à un déterminant défini préexistant.

4) Matras (2000b : 102), dans un article de synthèse consacré aux démonstratifs, mentionne la co-occurrence possible, dans certaines variétés, d'un démonstratif et d'un article défini dans le même syntagme. Citant l'exemple (49) de Boretzky (1994a : 55), il propose de l'analyser éventuellement comme « *a kind of conventionalised afterthought* ». Cette notion n'est toutefois pas définie plus précisément. On pourrait la traduire par « ajout après coup », « réflexion après coup », « pensée ultérieure », non planifiée cognitivement. Si l'*afterthought* est « conventionnalisé », il devient toutefois prévisible donc planifiable cognitivement. Selon Matras (2000b : 103-104), la postposition de l'adjectif rend obligatoire la présence d'un article défini l'accompagnant (DEM N DEF A et *DEM N A), de même que la simple postposition du démonstratif rend obligatoire la présence d'un article défini accompagnant le nom (DEF N DEM et *N DEM). La présence de l'article défini après le nom s'explique donc uniquement par la présence de l'adjectif après le nom, qui ne peut pas occuper seul cette position. On en déduit que la présence de l'article défini dépend de l'ordre des constituants nominaux. Le syntagme nominal s'arrête après le nom : « *One could hence argue that the postpositioned adjective falls outside the determiner scope of the demonstrative, and demands its own additional determination* » (Matras 2000b : 103).

5) Matras (2002 : 97) reprend et complète cette analyse : « *Adjectives in appositional function are generally treated as nominals, for case inflection as well as determination, and they may be introduced by a definite article, rendering the impression of a postposed definite adjective:* o rom o phuro *'ART man ART old = the old man'* ». Gaatone (2016 :16), reprenant le terme anglais, appelle « nominal » « tout mot, ou groupe de mots, susceptibles de fonctionner comme sujet de phrase, à savoir, un SUB avec ou sans déterminant (DET) » : un nom avec ou sans déterminant, un pronom, un infinitif, une proposition subordonnée. Il s'agit, ici d'une définition syntaxique.

Selon Matras (2002), l'antéposition est le seul ordre canonique pour un adjectif épithète – la fonction de l'adjectif postposé en DEF A DEF N n'est pas épithète (postposé), mais apposition. Qu'entend-on par « apposition », en quoi cette fonction diffère-t-elle d'épithète ou d'attribut ? D'une part, l'apposition relève de la catégories des « options » dans la chaîne syntagmatique du syntagme nominal de Matras (2002 : 166) (cf. section 1.1, p. 103). D'autre part, l'adjectif apposé est traité comme un nom, ce qui se manifeste par l'utilisation de désinences issues du paradigme nominal (exemple (52) de Boretzky 1993 : 171) et par l'utilisation d'un article défini qui l'accompagne (exemple (53) de Boretzky 1993 : 171). L'auteur réfute donc l'autre option théorique, celle d'un « adjectif postposé défini », c'est-à-dire semblant s'accorder avec le nom en genre, nombre, cas et définitude – comme c'est le cas dans les langues sémitiques. Cela appelle également une réflexion

sur ce que l'on entend par nom : ce qui est à droite d'un article défini est-il nécessairement de nature nominale ? doit-on ainsi supposer qu'un nom ne peut être la tête que d'un seul article défini ? Matras (2002 : 166) synthétise et théorise les analyses de Boretzky (1994a) et d'Igla (1996) :

> « *The postnominal 'option' slot deserves this designation due to the fact that adnominals that are accommodated here are often exempt from the constraints that apply to them in their usual, prenominal slot. Demonstratives are generally incompatible with definite articles. But when a demonstrative is postposed, then the noun it follows must be accompanied by a definite article: Kalderaš Vlax* o rom kadava *'this man'. Moreover, postposed demonstratives as well as postposed adjectives quite often carry nominal, rather than attributive, case agreement, reinforcing the impression that they serve as appositions: Kalderaš* e gažeskə kodoleskə *'for that man' (Boretzky 1994: 55). Igla (1996: 166) cites, from Agia Varvara Vlax, reduplication of the definite article with postposed adjectives, but not with postposed possessives:* i čhej i bari *'the big girl',* o dad tumaro *'your father'. While in some varieties such usages may be frequent, it seems that on the whole they are by far outnumbered by the conventional prenominal positioning of all attributes.* » (Matras 2002 : 166)

L'auteur met sur le même plan la postposition de différents éléments, démonstratif, adjectif et possessif, habituellement antéposés au nom. L'emplacement postnominal, celui des « options », qui a la fonction d'apposition, leur confère un statut intermédiaire entre leur nature initiale et une nature nominale. Il les contraint d'une part à porter une flexion de nom plutôt que d'adjectif ou de déterminant. Il provoque d'autre part l'émergence d'un article défini supplémentaire – qui accompagne le nom, dans le cas des démonstratif et possessif postposés, et qui accompagne l'adjectif, dans le cas de l'adjectif postposé. Doit-on en déduire que l'adjectif ne peut donc être placé « simplement » en position épithète postnominale (*DEF N A) – contrairement au démonstratif (DEF N DEM) et au possessif (DEF N POSS) ? L'ajout d'un article défini ne se fait en revanche qu'à l'adjectif : DEF N DEF A, mais *DEF N DEF POSS et *DEF N DEF DEM. L'auteur précise que l'usage de l'emplacement postnominal est, facultatif, peu fréquent et restreint à certaines variétés.

6) Le long article de Plank (2003b) consacré à la « double articulation » traite du romani à partir des données en **kalderaš de Suède** (Gjerdman et Ljungberg 1963). Plank (2003b : 344) compare **romani *vlax***[1] et grec : ils n'emploient qu'un seul article défini lorsque l'adjectif précède le nom, mais pratiquent la « ré-articulation lorsque l'adjectif est déplacé après le nom par emphase » (exemple (55). Il compare également le possessif et le démonstratif du romani : antéposés, l'un ne cohabite pas avec l'article défini (possessif : POSS N donc *POSS <u>DEF</u> N, *<u>DEF</u> N POSS), au contraire de l'autre (démonstratif : DEM N ou DEM <u>DEF</u> N) ; postposés, ils cohabitent tous deux avec un article défini précédant le nom (DEF N DEF POSS, exemple (56), et DEF N DEF DEM, exemple (57)).

(55) Romani (Plank 2003b : 344)

a.	*le*	*baré*	*raklés*	
	the	big	boy (masculine oblique)	
b.	*le*	*raklés*	*le*	*barés*
	the	boy	the	big

[1] Plank (2003b : 387) précise en note que, si toutes les variétés du romani non-*vlax* ont été en contact étroit avec le grec, toutes ne présentent pas cette structure. Il sous-entend que la structure pourrait être d'origine grecque, et que les variétés *vlax* l'auraient mieux conservée que les autres.

(56) Romani (Plank 2003b : 344-345)

a.	*murré*	*raklés*		
	my	boy (masculine oblique)		
b.	*le*	*raklés*	*le*	*murrés*
	the	boy	the	my

(57) Romani (Plank 2003b : 345)

a.	*kodolé*	*(le)*	*raklés*	
	that	(the)	boy (masculine oblique)	
b.	*le*	*raklés*	*le*	*kodolés*
	the	boy	the	that

(58) Romani (Plank 2003b : 344-345)

a.	*kodolé*	*murré*	*raklés*			
	that	my	boy (masculine oblique)			
b.	*le*	*raklés*	*le*	*kodolés*	*le*	*murrés*
	the	boy	the	that	the	my

Son apport est triple : d'une part il met d'emblée sur le même plan la postposition de l'adjectif, du possessif et du démonstratif, dans la mesure où ceux-ci peuvent être chacun *articulés* – ce qui suppose de les traiter comme des éléments autonomes. Le même principe serait est à l'œuvre dans la postposition, quelle que soit la nature de l'élément postposé. D'autre part, il est le premier à signaler l'existence des structures DEF N DEF POSS (56)b et DEF N DEF DEM (exemple (57)b) où les déterminants sont traités exactement comme des adjectifs. Enfin, il est le premier à montrer que le phénomène est récursif (Plank 2003b : 345) avec l'exemple (58) : DEM POSS N (antéposition), d'où DEF N DEF DEM DEF POSS (postposition et ré-articulation). On peut se demander si une structure DEF N DEF DEM DEF A serait possible. On peut en revanche s'étonner de ce qu'il ne mentionne pas la possibilité de postposer ces éléments sans ré-articulation : DEF N A (exemple (46), DEF N DEM (*e balen kodolen*, ces cochons, Boretzky 1994a : 55), DEF N POSS (*i dej miní*, ma mère, Igla 1996 : 165) qui sont pourtant attestés.

De même que Boretzky (2000b : 42-43) parlait de focus, Plank (2003b : 344-345) explique la ré-articulation par l'« emphase » sur l'adjectif. Il observe lui aussi que la forme des éléments n'est pas la même avant et après le nom : ils suivent une flexion de type adjectival lorsqu'ils sont antéposés, de type nominal lorsqu'ils sont postposés. On note cependant que le démonstratif, contrairement au possessif et à l'adjectif, ne suit pas une flexion adjectivale, mais une flexion propre au démonstratif. Plank (2003b : 345) ne considère pas que ces éléments revêtent une fonction d'apposition lâche ou d'*afterthought* disjoint, qui confère une plus grande autonomie à ces éléments. Il s'agit toujours d'épithètes, qui ont cependant « un plus haut degré de nominalité » que « leurs contreparties prénominales » (non ré-articulées) : en cela il rejoint l'interprétation de Matras (2002 : 97), en ce que ces éléments ont un statut nominal lorsqu'ils sont postposés.

7) Quelques variétés semblent contredire certaines de ces affirmations. C'est le cas du romani ***vlax* de Bogota**. Deman (2005 : 73) signale qu'un article défini « supplémentaire » peut émerger « entre le substantif et un adjectif postposé » : *o gras o parno*, le cheval blanc. Deman (2005 : 77) précise que, lorsque l'adjectif a une fonction épithète, il est habituellement postposé ; un article est alors inséré entre l'adjectif et le nom (exemple (59) :

(59) Romani (Deman 2005 : 77)

sik-av-el	*les*	*p-o*	*drom*	*o*	*lash-o*
zeig-TR-3SG	3SG.OBL.M	auf-SG.M	Weg	ART.SG.M	gut-SG.M

« er zeigt ihm (auf) den guten Weg »

En cas d'emphase, l'adjectif peut également être antéposé :

(60) Romani (Deman 2005 : 77)

sak-o	*bors*	*nev-o*	*s-as*	*jek*	*bar-o*	*krechuno*
jed-M	Jahr	neu-SG.M	Kop-IPF	INDF	groß-SG.M	Fest

« Jedes Neujahr war ein großes Fest »

On en déduit que, dans cette variété, la place non-marquée d'un adjectif épithète est postposée et articulée (DEF N DEF A), tandis que la place marquée, en cas d'emphase, est antéposée (DEF/IND A N). Ceci va à l'encontre des autres descriptions du romani, même balkanique (section 1.4.3, p. 138).

8) C'est également le cas de la variété de **Parakalamos**. Matras (2004 : 74-75) expose la flexion de type nominal des adjectifs épithètes et propose l'exemple suivant :

(61) Romani (Matras 2004 : 75)

kam-ama	*e*	*parn-es*	*e*	*grast-es*
vouloir-1SG.PRES	ART.DEF.M.SG.OBL	blanc-M.SG.OBL	ART.DEF.M.SG.OBL	cheval-M.SG.OBL

« Je veux le cheval blanc. »

L'auteur mentionne d'autres expressions : *o aver o dives*, le jour suivant (Matras 2004 : 93), *o terne o rakle*, les jeunes gens non-rom (Matras 2004 : 97). Il précise que l'article défini « *may be repeated to introduce each lexical member of the noun phrase* » – sans préciser plus avant pourquoi ou dans quel contexte. Il s'agit de structures polydéfinies à adjectif antéposé, DEF A DEF N – hapax dans les études sur le romani. Cette variété est-elle la seule à présenter cette construction ?

2.2. En grec

2.2.1. Similarité du grec et du romani

Le marquage multiple de la définitude existe dans plusieurs langues du monde et a fait l'objet d'analyses, notamment en grec moderne et dans les langues scandinaves (danois, suédois, norvégien). Nous présentons ici les travaux portant sur le grec moderne en particulier, ce pour plusieurs raisons :

- Le phénomène en grec présente une structure très similaire à celui du romani : réplication de l'article défini, qui est un morphème libre, ordre des mots libre, absence de poly*in*définitude, contextes d'emploi similaire…
- Il a été très étudié pour cette langue.

- Le romani a fait de nombreux emprunts structurels au grec, très tôt dans son histoire (Matras 2002 : 196-199). Le double marquage de la définitude pourrait être l'un d'entre eux.
- Comprendre celui du grec pourrait aider à comprendre celui du romani, dans ses similarités et ses divergences.
- Certaines questions n'ont pas été posées pour le grec. Nous allons les poser pour le romani et tenter d'y répondre. Cela pourra éclairer le phénomène grec en retour.

En grec moderne, l'adjectif comme le nom peuvent être précédés d'un article défini. Comme le romani, une variation syntaxique est possible en présence de l'adjectif : l'article vient se placer avant le nom et avant l'adjectif qui modifie celui-ci en (62)a et b. La structure est optionnelle : la grammaticalité de (62)c montre que cette répétition n'est pas obligatoire, même en présence d'un adjectif. La comparaison entre (62)b et d montre qu'en l'absence d'un tel article supplémentaire, l'adjectif ne peut venir se placer après le nom.

(62) Grec (Kolliakou 2004 : 264)

a.	*to*	*kokino*	*to*	*podilato*
	the-NEUT.SG	red-NEUT.SG	the-NEUT.SG	bike-NEUT.SG
b.	*to*	*podilato*	*to*	*kokino*
c.	*to*	*kokino*	*podilato*	
	the-NEUT.SG	red-NEUT.SG	bike-NEUT.SG	
d.	**to*	*podilato*	*kokino*	
	'the red bike'			

Les travaux de Kolliakou (1998) et Kolliakou (2003) qualifient (62)c de stratégie de modification de style **anglais** : l'adjectif précède le nom, l'article défini est unique et situé à la périphérie gauche du syntagme. Elle lui donne le nom de **monadique**. Elle qualifie (62)a et b de stratégie de modification de style **sémitique** : l'adjectif suit le nom et le syntagme comporte deux articles définis, précédant nom et adjectif. Elle lui donne le nom de **polydéfini**, pour souligner la présence multiple d'articles définis (autant d'articles qu'il y a d'adjectifs).

Afin de conserver le parallèle entre les deux structures (à un, deux articles ou plus) et d'opposer le syntagme présentant une marque unique de définitude à celui en présentant plusieurs, nous préférons nommer (62)c syntagme **<u>mono</u>défini**, (62)a et b syntagmes **<u>poly</u>définis**. Le marquage multiple de la définitude par réplication de l'article défini porte d'autres noms dans les recherches : *Definiteness Spreading* ou *Determiner Spreading* (Androutsopoulou 1995), « expansion de la définitude / du déterminant », plus précisément expansion aux différents adjectifs qui modifient le nom, et ce, autant de fois qu'il y a d'adjectifs. La double articulation en grec est présentée par Plank (2003b : 341-342, 344) et Alexiadou (2014 : 15-52), et comparée à celle des autres langues du monde, notamment d'Europe.

Notons enfin qu'en grec comme en romani, l'adjectif peut être utilisé sans le nom :

(63) Grec (Kyriakaki 2011 : 9)

a.	*θa*	*paris*	*to*	*makri*	*to*	*forema*	*i*	*to*	*kodo?*
	FUT	take-2SG	the-NEU	long-NEU	the-NEU	dress-NEU	or	the-NEU	short-NEU

'Are you getting the long dress or the short one?'

b. *To* *makri.*
the-NEU long-NEU
'The long one.'

2.2.2. Les constituants du syntagme polydéfini en grec

2.2.2.1. Nom

La question du type de nom pouvant entrer dans un syntagme polydéfini n'a pas été posée pour le grec. Nous essaierons, pour le romani, de dresser une typologie des noms rencontrés dans notre corpus, en suivant des catégories sémantiques et morphologiques, par exemple :

- noms propres / noms communs
- noms abstraits / noms inanimés / noms animés non-humains / noms animés humains
- noms indénombrables / noms dénombrables
- noms hérités / noms empruntés intégrés / noms empruntés non intégrés

Selon (Campos et Stavrou 2004 : 147), le nom n'est jamais accentué dans un syntagme polydéfini. Au contraire, dans un syntagme monodéfini, c'est le nom qui porte l'accent nucléaire. Il arrive cependant que l'adjectif porte l'accent si deux référents sont possibles pour le nom, et que l'adjectif reçoive un focus étroit servant à l'identification du référent. L'accentuation du syntagme polydéfini est donc uniforme (seul l'adjectif est accentué), tandis que celle du syntagme monodéfini dépend donc du contexte, puisque le nom comme l'adjectif peuvent porter l'accent.

Dans le syntagme polydéfini, lorsque l'adjectif suit le nom, il reçoit l'accent nucléaire. Il peut recevoir un accent contrastif et emphatique, si l'identification du référent le requiert. Lorsque l'adjectif précède le nom, il ne peut être qu'accentué contrastivement et présenter un focus étroit.

2.2.2.2. Adjectif

Du point de vue morphologique, les études ne précisent pas quel type d'adjectif est susceptible d'entrer dans une structure polydéfinie. Du point de vue syntaxique, les adjectifs qui ne peuvent pas remplir la fonction d'attribut du sujet ne peuvent entrer dans un syntagme polydéfini (Alexiadou 2001 : 107).

(64) Grec (Alexiadou 2001 : 107)

a. **i* *apofasi* *ine* *amerikaniki*
the decision is American

b.	**i*	*poli*	*amerikaniki*	*apofasi*
	the	very	American	decision
c.	**i*	*apofasi*	*i*	*amerikani*
	the	decision	the	American

Du point de vue sémantique, la plupart des chercheurs s'accordent à dire que seuls les adjectifs qui peuvent s'interpréter de façon intersective et restrictive peuvent entrer dans un syntagme polydéfini (Lekakou et Karatsareas 2016 : 191). Cette question est traitée plus avant en section 2.4.2, p. 179.

On sait que l'adjectif grec peut avoir son propre complément dans un syntagme monodéfini (exemple (65)a, b, c). Qu'en est-il dans un syntagme polydéfini ?

(65) Grec (Cinque 2010 : 47-48)

a.	*i*	*[perifani*	*ja*	*to*	*jo*	*tis]*	*mitera*
	the	proud	of	the	son	her	mother
	'the mother proud of her son'						
b.	*ta*	*prósfata*	*sideroména*	*me*	*prosohi*	*pukámisa*	
	the	recently	ironed	with	care	shirts	
	'the shirts recently ironed with care'						
c.	*o*	*kírios*	*kata*	*protereótita*	*lógos*		
	the	main	by	priority	reason		
	'the main reason in termes of priority'						

2.2.2.3. Déterminant

Avec quels articles peut se construire la structure polydéfinie ? La question a été largement étudiée pour les articles, mais peu pour les autres déterminants, tels que le démonstratif ou le possessif.

Article défini

C'est l'article défini qui fait l'objet d'une réplication. Le marquage multiple de la définitude peut se faire autant de fois qu'il y a d'adjectifs – c'est entre autres pourquoi on considère qu'il s'agit d'un type de modification adjectivale (Marinis 2003 : 165-166), et pourquoi certains chercheurs parlent d'expansion de la définitude ou du déterminant. On sait qu'elle est optionnelle : la présence de deux articles pour définir un seul nom, même modifié par un adjectif, n'est pas obligatoire. « *Polydefinites co-exist in the language with monadics* […]*, i.e. modification structures where only one determiner is present—although polydefinites have special semantic and pragmatic properties* […]. » (Lekakou et Szendrői 2007 : 129). En romani non plus, la structure qui nous intéresse n'est pas obligatoire.

Quelles sont les propriétés sémantiques et pragmatiques du syntagme polydéfini en romani ? Quelle est sa fréquence d'utilisation dans la langue ?

Le statut de l'article défini répliqué a été également examiné. On peut se demander si les deux articles ont le même statut ou si l'un (celui qui accompagne l'adjectif, parce que c'est le modifieur ?) est seulement une émanation de l'autre (celui qui accompagne le nom, parce que c'est la tête ?), ou bien s'il réalise une tête fonctionnelle distincte de l'article (Androutsopoulou 1995, Kolliakou 1998, Kolliakou 2004). Portent-ils tous deux le trait de la définitude ? Les théories suivantes ont été développées.

1) Androutsopoulou (1995), la première à avoir discuté le phénomène, et à l'avoir nommé *Definiteness Spreading*, a affirmé que chacun des articles représentait une tête distinctive dans le DP, abritant notamment les traits de définitude et de cas. La répétition de l'article défini consiste en la répétition du trait de définitude, qui est purement syntaxique et non sémantique. Le trait [+def], qui se manifeste par l'article défini, et qui est répété, n'est pas un trait porteur de sens. Sa répétition n'amène donc pas d'ajout de sens, pas de sur-détermination au sens propre : un tel syntagme n'est pas plus défini qu'un autre. Selon elle, l'exemple n'est possible qu'avec une interprétation prédicative sous-jacente, faisant de l'adjectif un attribut du sujet au sein d'une proposition relative sous-entendue.

(66) Grec (Lekakou et Szendrői 2007 : 146-147)

i	*olandiki*	*i*	*isvoli*	*mas*	*kseklirise*
the	Dutch	the	invasion	us	wiped out

'The dutch invasion wiped us out.'

Elle ne l'interprète pas comme un adjectif épithète. Cependant on a vu que la prédicativité n'était pas forcément un critère pour pouvoir entrer dans une structure polydéfinie. Comment expliquer la présence possible mais non obligatoire des articles définis antéposés au nom dans ces exemples (Marinis 2003 : 173) ?

2) Leu (2009) propose une analyse fondée sur la comparaison du grec avec les langues germaniques (allemande et scandinaves), où l'adjectif présente un paradigme flexionnel différent en fonction de la définitude du syntagme (cf. exemples (67)a et (68)a *versus* (67)a (68)b). Il postule qu'en grec, tous les adjectifs seraient accompagnés, par défaut, d'un article défini – comme le nom. Mais l'article de l'adjectif serait un reflet du complémenteur de la proposition relative sous-jacente au syntagme adjectival – il serait donc de nature différente de l'article défini qui accompagne le nom. L'optionnalité de la polydéfinitude s'explique, structurellement, par le fait que l'article accompagnant le nom est exprimé (polydéfinitude) ou non (monodéfinitude) dans la forme de surface (exemples (69)a et (69)b).

(67) Allemand (Leu 2009 : 3-4)

a.	*d-er*	*schön-e*	*Tisch*
	the-AGRA[1]	pretty-WK	table

[1] « *I am glossing strong inflection as* AGRA, *mnemonic for "adjectival agreement*. » (Leu 2009 : 3)

b.	*ein*	*schön-er*	*Tisch*
	a	pretty-AGRA	table

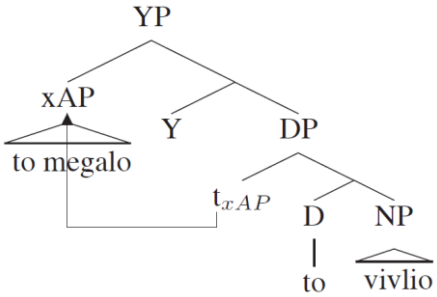
(68) Suédois (Leu 2009 : 4)

a.	*de-t*	*stora*	*hus-et*
	the-AGRA	big	house-DET
b.	ett	*stor-t*	*hus*
	a	big-AGRA	house

(69) Grec (Leu 2009 : 13-14)

a.	*to*	*megalo*	*to*	*vivlio*
	the	big	the	book

b.



3) L'article défini suppose un principe d'existence et d'unicité, or ce principe est violé.

(70) Grec (Lekakou et Szendrői 2014 : 219, inspiré de Kolliakou 2004 : 269)

a. Speaker A:	*Ti*	*pires*	*tu*	*Janni*	*ja*	*ta*	*christujena ?*
	what	took.2SG	the.GEN	John	for	the	Christmas

'What did you get John for Christmas?'

b. Speaker B:	*(Tu*	*pira)*	*tin*	*asimenia*	*pena.*
	him.GEN	took.1SG	the	silver	pen

'(I got him) the silver pen.'

c. Speaker A:	*Ti*	*pires*	*tis*	*Marias ?*
	what	took.2SG	the.GEN	Maria

'What did you get for Maria?'

d. Speaker B:	*(Tis*	*pira)*	*tin*	*pena*	*ti*	*chrisi.*
	her.GEN	took.1SG	the	pen	the	golden

'(I got her) the golden pen.'

La présence de deux articles définis en (70)d n'implique pas l'existence de deux stylos, ni de deux stylos dorés. Il n'est pas dit pour autant qu'il n'y ait qu'une seule entité dorée. En revanche, il y a bien, dans le contexte, deux stylos. Deux articles définis sont employés pour renvoyer à un seul référent. C'est que l'interprétation du syntagme doit se faire de façon globale. L'unicité vaut pour l'ensemble du syntagme, car il représente l'intersection de l'ensemble des stylos et de l'ensemble des entités dorées. Il y a donc dans la structure deux déterminants, mais un seul référent, c'est-à-dire un seul nom qui corresponde à un objet du monde. La présupposition d'unicité associée à la définitude vaut, selon Lekakou et Szendrői (2012 : 145), pour le grand DP commun et complexe, et non pour ses sous-parties individuelles, qui n'ont pas de contribution sémantique individuelle à la définitude.

4) Lekakou et Szendrői (2012 : 145) montrent qu'aucun des deux articles n'est « réel » d'un point de vue sémantique : « *Thus, it seems that, semantically, both definite determiners are vacuous at the position where they surface. That is not to say, however, that there is no definiteness in the*

polydefinite: what is unique is the element in the intersection of the two DP categories. In other words, definiteness is interpreted on, or actually above, the larger DP. » Nous reviendrons plus loin sur l'interprétation restrictive du syntagme, mais l'on peut déjà noter l'idée de vacuité des deux articles. Ils viendraient remplir une position sans porter le trait de définitude qui leur est habituellement associé. Les deux autrices postulent en revanche l'égalité entre eux : ils ont les mêmes propriétés sémantiques et syntaxiques, et les mêmes propriétés que l'article d'un syntagme monodéfini. Les deux sous-parties du syntagme polydéfini sont donc des DP complets (Lekakou et Szendrői 2007, Lekakou et Szendrői 2012). Cette approche réfute celle de la proposition relative sous-jacente, pour laquelle la sous-partie « adjectivale » du syntagme polydéfini entre dans une relation prédicative avec la sous-partie « nominale ». Cette approche considère que les deux sous-parties jouent, ensemble, un rôle dans la référence et que la syntaxe sous-jacente est symétrique, ce qui explique les phénomènes d'accord. Selon Lekakou et Szendrői (2014 : 213), supposer que les articles d'un syntagme polydéfinis sont identiques, explique les phénomènes d'accord, la morphologie identique des articles : ce sont tous les instances d'un même élément.

5) Lekakou et Szendrői (2012) examinent successivement deux approches de l'article défini. La première propose que l'article défini soit explétif. La définitude n'est alors pas encodée, en grec, par l'article défini, mais par un opérateur phonologiquement nul, parfois reflété par l'article défini dans la structure de surface, ou par plusieurs, ou par aucun – que ce soit dans les syntagmes monodéfinis ou polydéfinis (le premier ou le deuxième article). L'assertion et la présupposition d'existence et d'unicité, n'est pas réalisée, en grec, par l'article défini. À vrai dire, l'article défini ne contribue jamais à la définitude, du point de vue sémantique (Lekakou et Szendrői 2014 : 214). Les noms propres sont toujours accompagnés d'un article défini et par conséquent conçus comme définis par la grammaire traditionnelle (Lekakou et Karatsareas 2016 : 199). DefP serait placé au-dessus de DP dans la hiérarchie de l'arbre, et toujours rempli par ∅. La présence d'un ou de deux articles définis comme tête de DP en-dessous n'a donc aucune espèce d'impact sur la définitude du syntagme. Cela signifie alors que la polydéfinitude porte mal son nom : un syntagme présentant plusieurs articles définis ne serait pas à proprement parler *polydéfini*, et en tout cas plus défini qu'un syntagme en présentant un seul. Une autre preuve de la vacuité sémantique de l'article défini du grec est qu'il accompagne de manière obligatoire les noms propres. L'article défini est là sémantiquement explétif, il ne marque pas la définitude. Les autrices parlent d'une scission entre les deux têtes Def et D (*Def-D split*) c'est-à-dire entre la définitude et le déterminant (ici, l'article défini). Le lieu de la définitude sémantique est, selon Lekakou et Szendrői (2012 : 138-139), Def et non D, aussi bien dans les syntagme mono- que polydéfinis.[1] Cette théorie présente donc le même type d'analyse en termes de définitude, pour les syntagmes mono- et polydéfinis (Lekakou et Karatsareas 2016 : 235). Il n'y a donc pas d'ambiguïté lexicale entre les articles définis d'un syntagme polydéfini. La scission entre Def et D serait due à l'existence du cas dans la langue, plus exactement au fait que le plus haut élément de la projection nominale étendue soit *Kase*.

La deuxième analyse propose que l'article défini encode bel et bien la définitude, mais que ses deux exemplaires, dans le syntagme polydéfini, n'aient pas le même statut : l'un serait « réel » et

[1] Dans les langues du monde sans article défini, on aurait ainsi un DefP avec une tête nulle. Cela expliquerait que les syntagmes nominaux y soient interprétés « par défaut » comme définis.

l'autre « sémantiquement explétif », sans contribution sémantique à la définitude du syntagme. Autre possibilité : l'un serait « réel », et l'autre serait la réalisation d'une tête syntactique distincte et différente. Reste à fixer si c'est celui qui accompagne le nom ou celui qui accompagne l'adjectif – des arguments existent dans les deux sens. Que faire en outre des syntagmes présentant plusieurs adjectifs et donc plusieurs articles : peut-il y avoir plusieurs articles « réels » ? Comment décider lequel il est ? Si l'article défini est toujours ou parfois explétif, il serait seul à fonctionner ainsi dans la langue. Il n'y a donc aucune motivation, aucun autre cas similaire dans la langue, qui puisse venir étayer cette hypothèse – contrairement à la première analyse, qui est étayée par le fonctionnement des noms propres.

6) Panagiotidis (2000) considère l'article accompagnant le nom comme un article explétif. La forme de l'article défini et de l'article explétif est selon lui identique en grec standard : au masculin singulier, *o* dans les deux cas. Ce n'est pas le cas dans la variété septentrionale du grec, où, toujours au masculin singulier, l'article défini est *u* tandis que l'article explétif est *i*. Il donne l'exemple du catalan et du frison de Föhr qui présente une même distinction entre les articles. C'est pour lui une preuve qu'il existe, en grec standard également, deux entrées lexicales différentes pour l'article (dans le dictionnaire mental des locuteurs) : un article à proprement parler défini, et un article explétif, qualifié d'*expletive place-holder*. La forme explétive ne possède pas, selon lui, le trait de la définitude, contrairement à l'article défini proprement dit (Panagiotidis 2000 : 731). Une autre façon de les distinguer est la possibilité, pour l'article défini mais non pour l'article explétif, de co-occurrer avec le démonstratif (Panagiotidis 2000 : 731-732).

7) Kyriakaki (2011 : 58) critique l'idée que l'article défini soit explétif. Selon elle, le fait que les noms propres soient toujours accompagnés d'un article défini ne constitue pas une preuve. Les noms propres et les noms communs dénombrables peuvent être employés nus, au vocatif et en fonction attribut. Elle montre que l'article défini a un apport sémantique. Il marque bel et bien la définitude, mais il est sous-spécifié. C'est pourquoi il s'emploie avec les noms propres (Kyriakaki 2011 : 58-59). L'article défini est non pas explétif, mais sous-spécifié : « *It is underspecified, but context, by means of the uniqueness operator, fills out the specification. It is correctly predicted then that in languages which have polydefinites, the determiner will be underspecified in terms of definiteness, and vice versa* » (Kyriakaki 2011 : 221). En revanche, la comparaison avec l'anglais n'est pas pertinente. En anglais, le modifieur nominal en apposition étroite (DEF N <u>DEF N</u>) dénote une entité unique, à l'instar d'un DP ayant la fonction d'attribut du sujet dans une proposition verbale. Ce n'est pas le cas en grec, où l'article défini ne sélectionne pas une entité unique. C'est pourquoi la présence de plusieurs articles définis dans le syntagme n'est pas un problème, et contribue à l'identification du référent. Si **My friend the doctor the Italian* est agrammatical en anglais, ce n'est pas le cas de son équivalent *grec o filos mu o daskalos o Italos*, mon ami le professeur italien (littéralement mon ami le professeur l'italien) (Kyriakaki 2011 : 58). En grec, la définitude comprend deux composantes : la familiarité et l'unicité, qui sont encodées dans deux projections différentes (Kyriakaki 2011 : 58). Or l'article défini du grec n'encode que la familiarité, c'est-à-dire le fait que le référent soit connu, ait déjà été mentionné. C'est pourquoi les noms communs, noms propres et syntagmes génitivaux font l'objet de modifieurs nominaux restrictifs, ou apposition étroite.

Flexion

En grec, les noms se fléchissent en nombre et en cas. Le genre est fixé lexicalement. Ils appartiennent à différentes classes flexionnelles, qui déterminent leur désinence. Les adjectifs se fléchissent eux aussi en nombre, en cas et en genre, en accord avec le nombre, cas et genre du nom modifié. Ils appartiennent eux aussi à différentes classes flexionnelles, indépendantes de la classe du nom. L'adjectif s'accorde avec le nom, quelle que soit sa fonction, épithète ou attribut. Le déterminant se fléchit en nombre, cas et genre et s'accorde avec le nom (Alexiadou, Horrocks et Stavrou 1999 : 7-8). Si le syntagme est fléchi, les marques de cette flexion sont présentes sur l'adjectif, quelle que soit sa position, mais également sur les articles définis (Lekakou et Szendrői 2007) :

(71) Grec (Lekakou et Szendrői 2007 : 139)

[$_{KP}$	*K*	*[$_{DP}$*	*DP*	*tu*	*palju]*	*[$_{DP}$*	*tu*	*spitiu]]]*
				the-GEN	old-GEN		the-GEN	house-GEN

Dans l'exemple (71), la marque de génitif présente sur l'article défini et le nom dans la partie « nominale » du syntagme polydéfini se trouve également sur l'article défini et l'adjectif dans la partie « adjectivale ». Si le syntagme est introduit par une préposition, celle-ci peut être répétée avant le deuxième article :

(72) Grec (Lekakou et Szendrői 2007 : 139)

a.		*[$_{PP}$*	*P*	*me*	*[$_{KP}$*	*K*	*[$_{DP}$*	*[$_{DP}$*	*to*	*kokino]*	*[$_{DP}$*	*to*	*podhilato]]]]*			
				with					*the-ACC*	*red-ACC*		*the-ACC*	*bicycle-ACC*			
b.	*	*[$_{PP}$*	*P*	*me*	*[$_{KP}$*	*K*	*[$_{DP}$*	*[$_{DP}$*	*to*	*kokino]*	*[$_{PP}$*	*P*	*me*	*[$_{DP}$*	*to*	*podhilato]]]]]*
				with					*the-ACC*	*red-ACC*			*with*		*the-ACC*	*bicycle-ACC*

Dans l'exemple (72), la préposition qui introduit la partie « nominale » introduit également la partie « adjectivale ». Chaque sous-partie du syntagme polydéfini constitue un DP sémantiquement et syntactiquement entier, à la construction symétrique. Seul manque le nom dans la partie « adjectivale » du syntagme polydéfini. Les deux DP forment un grand DP commun.

Si l'on considère qu'il y a plutôt *spreading*, expansion, ou encore réplication, émanation, copie, de l'article défini ou de la définitude, il nous faut postuler la réplication de ces trois éléments : article défini, désinence casuelle et préposition.

Démonstratif

En grec, le démonstratif est un déterminant qui ne s'emploie jamais seul, mais combiné à l'article défini (DEM DEF N ou DEF N DEM). Lekakou et Szendrői (2014) ne précisent pas si le démonstratif peut se trouver dans un syntagme polydéfini. Kyriakaki (2011 : 21-23) précise qu'il le peut, en position initiale de syntagme (exemples (73)a, (73)b), en position clitique du nom (exemple (73)c), ce sont deux positions habituelles pour lui. Les syntagmes de l'exemple (73) sont triplement déterminés, par deux articles définis et un démonstratif.

(73) Grec (Kyriakaki 2011 : 21-22)

a.	DEM	DET	ADJ	DET	N
	afti	*i*	*asimenja*	*i*	*pena*
	this	the	silver	the	pen
b.	DEM	DET	N	DET	ADJ
	afti	*i*	*pena*	*i*	*asimenja*
	this	the	pen	the	silver
c.	DET	N	DEM	DET	ADJ
	i	*pena*	*afti*	*i*	*asimenja*
	the	pen	this	the	silver
d.	DET	ADJ	DEM	DET	N
	i	*asimenja*	*afti*	*i*	*pena*
	the	silver	this	the	pen

'this silver pen'

Mais il peut également être placé en position clitique de l'adjectif (exemple (73)d). Une telle position est marginale, et peut également se produire en syntagme monodéfini : DEF A DEM N (Kyriakaki 2011 : 22). Enfin, il peut se placer en fin de syntagme, en l'absence d'adjectif (DEF N DEM) et éventuellement en la présence d'un adjectif non articulé (DEF A N DEM), mais difficilement si l'adjectif est articulé (exemples (74)a et (74)b). Plus le syntagme est lourd, moins le démonstratif peut apparaître, ainsi (74)c qui contient deux adjectifs articulés (Kyriakaki 2011 : 23).

(74) Grec (Kyriakaki 2011 : 22-23)

a.	*/?	DET	N	DET	ADJ	DEM		
		i	*pena*	*i*	*asimenja*	*afti*		
		the	pen	the	silver	this		
b.	*/?	DET	ADJ	DET	N	DEM		
		i	*asimenja*	*i*	*pena*	*afti*		
		the	silver	the	pen	this		
c.	*	DET	ADJ	DET	ADJ	DET	N	DEM
		i	*asimenja*	*i*	*palja*	*i*	*pena*	*afti*
		the	silver	the	old	the	pen	this

Possessif

En grec, le possessif n'est pas un déterminant, mais un pronom clitique qui s'emploie avec l'article défini (DEF N POSS) (Alexiadou 2005 : 804). Lekakou et (Szendrői 2014) ne précisent pas si le possessif peut se trouver dans un syntagme polydéfini. Le possessif peut également intervenir dans une structure polydéfinie (Kyriakaki 2011 : 28-29). Il apparaît dans sa position habituelle de clitique du nom ((75)a) et, comme on l'a vu pour le démonstratif avec (73)d, il peut également être clitique de l'adjectif ((75)b) :

(75) Grec (Kyriakaki 2011 : 28-29)

a.	*to*	*vivlio*	*mu*	*to*	*paljo*
	the	book	cl.1s	the	old
b.	*to*	*paljo*	*mu*	*to*	*vivlio*
	the	old	cl.1s	the	book

'my old book'

Le possessif peut être clitique de l'adjectif, même en syntagme monodéfini (DEF A POSS N) (Kyriakaki 2011 : 28). Les syntagmes de l'exemple (75) sont triplement déterminés, par deux articles définis et un possessif. Il existe selon Kyriakaki (2011 : 109) une différence de sens subtile entre les deux ordres DEF A POSS DEF N (76) (76)a): et DEF A DEF N POSS (76)b): le premier met en valeur la possession ou l'information apportée par le déterminant possessif, tandis que le second met en valeur le couple adjectif-nom, le référent de ce couple plutôt que le lien de possession. Un syntagme monodéfini (DEF A N POSS) est, quant à lui, ambigu entre les deux.

(76) Grec (Kyriakaki 2011 : 108-109)

a.	*to*	*kokkino*	*mu*	*to*	*forema*
	the.N	red.N	my	the.N	dress.N
b.	*to*	*kokkino*	*to*	*forema*	*mu*
	the.N	red.N	the.N	dress.N	my

'my red dress'

Article indéfini

L'indéfinitude se marque, en grec, soit par l'usage de l'article indéfini, soit par une détermination zéro (Alexiadou, Horrocks et Stavrou 1999 : 7-8). Il n'existe pas, en grec, de poly*in*définitude, c'est-à-dire de réplication de l'article indéfini de type IND N IND A, ou IND A IND N (Lekakou et Szendrői 2007 : 138), (Lekakou et Szendrői 2012 : 123-124). Selon les autrices, la structure polydéfinie relève d'un phénomène général de *close apposition*, qui n'accepte pas d'article indéfini (« *close apposition necessarily involves two definite DPs.* […] *Since polydefinites are an instance of close apposition, the ban against indefinites is expected here too.* » (Lekakou et Szendrői 2007 : 144).

Si l'on voulait postuler l'existence de la poly*in*définitude, ce serait en postulant un article zéro : Ø N Ø A, ou Ø A Ø N (Lekakou et Szendrői 2007 : 18). Du point de vue formel, ce que l'on prend habituellement pour l'article indéfini serait alors le quantifieur numéral *un*, n'ayant pas été grammaticalisé en article indéfini. Le véritable article indéfini serait un article nul, réduplicqué lorsque l'adjectif épithète est postnominal (Cinque 2010 : 146). Postuler la réplication d'un morphème zéro ou d'une absence de morphème est toutefois difficile à démontrer.

Leu (2009 : 17) explique l'impossibilité de la polyindéfinitude au moyen d'arguments structurels générativistes. À la suite de Giusti (1995) et Stavrou (2012), Lekakou et Szendrői (2012 : 24) l'expliquent par l'absence d'un véritable article indéfini. On est induit en erreur par le numéral ou quantifieur « un » qui lui est formellement identique. Or le numéral ne peut pas jouer un rôle d'identification, contrairement à l'article défini. C'est pourquoi il peut se combiner avec lui[1] :

(77) Grec (Lekakou et Szendrői 2012 : 124)

O	*enas*	*drastis*	*sinelifthi.*
the	one	perpetrator	arrested.3SG.NONACT

'One perpetrator was arrested. '

[1] Il en va de même en romani (exemple (20).

D'un point de vue diachronique, on peut interpréter ceci comme une preuve que les numéraux ordinaux ένας (MASC), μία (FEM) et ένα (N) n'ont pas achevé le processus de grammaticalisation – or on ne réduplique pas les numéraux et les quantifieurs en grec. Alexiadou (2014 : 95-109), dans son chapitre consacré aux déterminants multiples dans les syntagmes indéfinis, partage le point de vue selon lequel ένας est en réalité un numéral. Elle étudie même l'hypothèse d'une expansion de l'indéfinitude via un article zéro, mais montre que c'est en fait structurellement impossible.

Alexiadou et Wilder (1998) ne partagent cependant pas l'avis général, et supposent qu'il y a bel et bien polydéfinitude dans un syntagme indéfini, ce que suggère la possibilité de postposer l'adjectif (IND N A, IND N A A, IND A N A). Avec un article défini, il faut une structure polydéfinie pour postposer l'adjectif. Cela doit donc fonctionner de même avec un article indéfini. La solution proposée est une structure polydéfinie avec article défini muet, la tête de DefP étant vide, en présence de l'indéfini. Alexiadou (2014 : 105-109) examine la possibilité structurelle de la polyindéfinitude en grec (*spreading of indefiniteness*).

Quantifieur

A notre connaissance, les études ne disent pas si un quantifieur peut entrer dans un syntagme polydéfini. Kyriakaki (2011 : 24-26, 193-194) mentionne όλος, *olos*, tout, qui peut se trouver en syntagme monodéfini, y compris en présence d'un démonstratif et d'un posssessif. Cela donne lieu à des syntagmes triplement déterminés, tels que QUANT DEM DET A N ou DEM DET A N QUANT. Mais elle ne donne pas d'exemple en syntagme polydéfini.

2.2.2.4. Ordre des constituants

Dans un syntagme **indéfini**, l'adjectif peut être antéposé (exemple (78)a) au nom, postposé (exemple (78)b, c, d) ou antéposé et postposé lorsque les adjectifs sont plusieurs (exemple (78)e, f).

(78) a. Grec (Alexiadou 2014 : 105)

ena	*megalo*	*kokkino*	*vivlio*
a	big	red	book

b. Grec (Alexiadou, Horrocks et Stavrou 1999 : 8)

mia	*jineka*	*eksipni*
a	woman	clever

c. Grec (Alexiadou 2014 : 106)

ena	*vivlio*	*megalo*	*kokkino*
a	book	big	red

d. Grec (Alexiadou 2014 : 106)

ena	*vivlio*	*kokkino*	*megalo*
a	book	red	big

e. Grec (Alexiadou 2014 : 105)

ena	*megalo*	*vivlio*	*kokkino*
a	big	book	red

f. Grec (Alexiadou 2014 : 105)

ena	*kokkino*	*vivlio*	*megalo*
a	red	book	big

Dans un syntagme défini **monodéfini**, l'adjectif est antéposé (exemple (79), qu'il y ait un ou plusieurs adjectifs (exemple (80).

(79) Grec (Lekakou et Szendrői 2012 : 108)

a.		*i*	*asimenia*	*pena*
		the	silver	pen
b.	*	*i*	*pena*	*asimenia*
		the	pen	silver

(80) Grec (Kyriakaki 2010 : 2)

a.		*i*	*asimeɲa*	*xiropijti*	*pena*
		the-NOM.FEM	silver-FEM	handmade-FEM	pen-FEM
		'the silver, handmade pen'			
b.	*	*i*	*asimeɲa*	*pena*	*xiropijti*
		the-NOM.FEM	silver-FEM	pen-FEM	handmade-FEM
		'The silver pen handmade'			

Dans un syntagme **polydéfini**, l'adjectif est soit antéposé (exemple (81)a), soit postposé (exemple (81)b). L'ordre des mots est donc plus libre qu'en syntagme monodéfini (Alexiadou et Wilder 1998 : 303-304, Cinque 2010 :105).

(81) Grec (Lekakou et Szendrői 2014 : 213)

a.	*i*	*asimenia*	*i*	*pena*
	the-FEM.NOM	silver	the-FEM.NOM	pen
b.	*i*	*pena*	*i*	*asimenia*
	the-FEM.NOM	pen	the-FEM.NOM	silver

Dans les langues du monde présentant une structure polydéfinie, l'adjectif est le plus souvent postposé (type DEF N DEF A). Le grec permet toutefois l'antéposition de l'adjectif (type DEF A DEF N) : cet ordre est « possiblement plus marginal que la variante postnominale » (Plank 2003b : 351). Existe-t-il une différence de sens entre les deux ? La postposition serait une façon spéciale d'emphatiser l'adjectif (Plank 2003b : 351). Stavrou (c.p.) citée par Cinque (2010 : 107) le confirme : « *with articulated adjectives, the postnominal order is the unmarked one, the prenominal being necessarily focused* ».

Selon Campos et Stavrou (2004 : 145-146), la structure DEF N DEF A assigne plutôt une propriété temporaire au référent du nom, tandis que la structure DEF A N transmet une propriété permanente, inhérente. Toutefois, la structure DEF A DEF N peut à la fois transmettre une propriété permanente (leur exemple 18 : *To kalo to palikari kseri ki alo monopati, The good guy can find his way*) et une propriété temporaire (leur exemple 20a : *I asterismi i orati perilamvanun ti Megali Arkto, The visible constellations include the Big Ursus*).

Cette différence affleure dans les cas où l'adjectif peut revêtir un sens propre et un sens figuré.

Alexiadou et Wilder (1998 : 314) parlent d'adjectif « ambigus », car ils ont développé un sens différent selon leur emploi épithète ou attribut. C'est le cas du grec *ftoxos*, qui signifie littéralement « pauvre, appauvri », mais aussi « pitoyable, malheureux » au sens figuré. Employé en fonction attribut, il ne peut revêtir que son sens propre. Antéposé dans un syntagme monodéfini, il revêt son sens propre ou bien son sens figuré (exemple a). Antéposé ou postposé dans un syntagme polydéfini, il ne peut revêtir que son sens propre selon Alexiadou et Wilder (1998 : 314-315) (exemple (82)b et c). Le syntagme polydéfini fonctionnerait donc comme l'attribut du sujet, sur le plan sémantique.

(82) Grec (Alexiadou et Wilder 1998 : 314-315)

a. *o* *ftoxos* *anthropos*
the poor man
'the impoverished/pitiable man'

b. *o* *ftoxos* *o* *anthropos*
the poor the man
'the impoverished/*pitiable man'

c. *o* *anthropos* *o* *ftoxos*
the man the poor
'the impoverished/*pitiable man'

Mais il y a débat puisque le sens figuré, non-intersectif, serait acceptable si l'adjectif est antéposé selon Stavrou (Cinque 2010 : 147). Dans ce cas, l'adjectif polydéfini changerait de sens selon sa position et selon le type de syntagme (monodéfini ou polydéfini) (Cinque 2010 : 107). L'adjectif non prédicatif *kaimenos*, de même sens, se comporte de manière similaire :

(83) Grec (Androutsopoulou 1995 : 24, Androutsopoulou 2001 : 191, Cinque 2010 : 107)

a. *o* *kaimenos* *o* *mathitis*
the pitiable the student
'the pitiable student'

b. **aftos* *o* *mathitis* *ine* *kaimenos*
this the student is pitible
'This student is pitible.'

Terzi (Cinque 2010 : 147) affirme que la structure polydéfinie est même obligatoire pour que cet adjectif prenne son sens figuré. Cela confirme selon lui le sens figuré de malheureux que peut prendre *ftohos* en syntagme polydéfini. Nous avons vu en section 3.5, p. 97, que l'on trouve en français la même ambiguïté de sens, et que c'est la position vis-à-vis du nom qui en décide : le sens propre en position antéposée et postposée, le sens figuré en position antéposée seulement.

2.2.2.5. Ordre des adjectifs

Lorsqu'il y a plusieurs adjectifs[1], l'antéposition et la postposition se combinent (84) et (85).

(84) Grec (Lekakou et Szendrői 2012 : 121)

[1] Selon Stavrou (Cinque 2010 : 107), un seul adjectif serait possible de façon « naturelle » dans un syntagme polydéfini.

a.	*i*	*pena*	*i*	*asimenia*	*i*	*kenurja*
	the	pen	the	silver	the	new
b.	*i*	*pena*	*i*	*kenurja*	*i*	*asimenia*
c.	*i*	*asimenia*	*i*	*pena*	*i*	*kenurja*
d.	*i*	*asimenia*	*i*	*kenurja*	*i*	*pena*

(85) Grec (Kyriakaki 2010 : 2)

a.	*I asimeɲa i xiropijti i pena*	The silver the handmade the pen	DEF A_1 DEF A_2 DEF N
b.	*I xiropijti i asimeɲa i pena*	The handmade the silver the pen	DEF A_2 DEF A_1 DEF N
c.	*I asimeɲa i pena i xiropijti*	The silver the pen the handmade	DEF A_1 DEF N DEF A_2
d.	*I xiropijti i pena i asimeɲa*	The handmade the pen the silver	DEF A_2 DEF N DEF A_1
e.	*I pena i xiropijti i asimeɲa*	The pen the handmade the silver	DEF N DEF A_2 DEF A_1
f.	*I pena i asimeɲa i xiropijti*	The pen the silver the handmade	DEF N DEF A_1 DEF A_2

La règle générale veut que les adjectifs articulés ne puissent pas être plus proches du nom que les adjectifs non articulés (Cinque 2010 : 105-106).

(86) Grec (Cinque 2010 : 105-106)

a.	*	*i*	*megali*	*kenuria*	*[i*	*kokkini]*	*valitsa*
		the	big	new	the	red	suitcase
		'the big new red suitcase'					
b.		*i*	*megali*	*i*	*kenuria*	*kokkini*	*valitsa*
		the	big	the	new	red	suitcase
		'the big new red suitcase'					

L'autre règle veut que tout adjectif postposé soit accompagné d'un article défini, ce qui n'est pas le cas de l'adjectif antéposé, qui peut partager le même article défini initial que le nom, ou bien avoir le sien propre (Marinis 2003) :

(87) Grec (Marinis 2003 : 167)

a.		*to*	*meghalo*	*petrino*	*spiti.*	
		the	big	stone-made	house	
b.	*	*to*	*spiti*	*meghalo*	*petrino.*	
		the	house	big	stone-made	
		'the big house, made of stone.'				
c.	*	*to*	*spiti*	*meghalo*	*to*	*petrino.*
		the	house	big	the	stone-made
d.	*	*to*	*spiti*	*to*	*meghalo*	*petrino.*
		the	house	the	big	stone-made
		'the big house, made of stone.'				
e.		*to*	*meghalo*	*to*	*petrino*	*spiti.*
		the	big	the	stone-made	house
f.		*to*	*meghalo*	*spiti*	*to*	*petrino.*
		the	big	house	the	stone-made
		'the big house, made of stone.'				

Les exemples de la série (87) diffèrent de la série (88) en ce que dans cette dernière, chaque sous-partie de syntagme est introduite par un article défini.

(88) Grec (Marinis 2003 : 166-167)

a.	*to*	*meghalo*	*to*	*petrino*	*to*	*spiti*
	the	big	the	stone-made	the	house
b.	*to*	*meghalo*	*to*	*spiti*	*to*	*petrino*
	the	big	the	house	the	stone-made
c.	*to*	*petrino*	*to*	*spiti*	*to*	*meghalo*
	the	stone-made	the	house	the	big
d.	*to*	*spiti*	*to*	*meghalo*	*to*	*petrino*
	the	house	the	big	the	stone-made
e.	*to*	*spiti*	*to*	*petrino*	*to*	*meghalo*
	the	house	the	stone-made	the	big

'the big house, made of stone.'

On observe la répétition systématique de l'article et l'ordre flexible des adjectifs. Les études ne s'accordent pas quant aux éventuelles nuances sémantiques ni aux contextes d'emploi de ces différents ordres. Cette flexibilité n'existe qu'aussi longtemps que l'article défini est répété. Dans le cas contraire, l'adjectif n'est qu'antéposé :

(89) Grec (Lekakou et Szendrői 2007 : 143)

a.	*	*to*	*spiti*	*megalo*	*to*	*petrino*
		the	house	big	the	stone
b.		*to*	*megalo*	*spiti*	*to*	*petrino*
		the	big	house	the	stone
c.	*	*to*	*spiti*	*to*	*megalo*	*petrino*
		the	house	the	big	stone
d.		*to*	*megalo*	*petrino*	*to*	*spiti*
		the	big	stone	the	house
e.	*	*to*	*petrino*	*megalo*	*to*	*spiti*
		the	stone	big	the	house

Notons qu'il est possible, en présence de deux adjectifs, de ne répéter l'article qu'une seule fois et non deux, si l'un des adjectifs est antéposé (89)b. Ce n'est pas le cas des adjectifs postposés, qui doivent tous être accompagnés d'un article (89)c. En réalité, la postposition de l'adjectif continue d'être contrainte en grec : elle n'est possible que si l'adjectif est accompagné d'un article défini. Les exemples (89)a et c montrent que tout adjectif postposé doit être accompagné de l'article défini. Un adjectif sans article défini doit précéder le nom (Lekakou et Szendrői 2007 : 143, Panagiotidis et Marinis 2011). Mais alors, l'exemple (89)e est étonnant, dans la mesure où les deux adjectifs sont justement antéposés, et où l'un d'entre eux devrait pouvoir être monadique. L'exemple (89)d est licite, alors qu'il est quasiment identique à l'exemple (89)e, à ceci près que les deux adjectifs antéposés sont inversés. Selon Lekakou et Szendrői (2007 : 143), l'inversion des adjectifs *megalo* et *petrino* n'est pas possible car *petrino* et *spiti* forment un ensemble qui ne peut être scindé. C'est ainsi que nous comprenons « *It seems to us to involve some sort of idiomatic, complex adjective* ». Cela semble indiquer que tous les adjectifs ne sont pas sur le même plan vis-à-vis du nom, et que l'un (*petrino*) ne peut pas être loin de lui. La « maison en pierre » est un type de maison dans la typologie des maisons : maison en bois, maison en paille, maison en brique… Cela renvoie à un mécanisme pour nous comparable à celui décrit pour un syntagme tel qu'*énergie nucléaire* (cf. section 2.4.3, p, 185). Rappelons également que l'ordre « universel » des adjectifs serait : évaluatif – taille – forme – couleur (cf. section 2.2.3, p. 77). Cela expliquerait que l'adjectif de matière *petrino* doive rester plus proche du nom que l'adjectif de taille *megalo*.

Selon Stavrou, citée par Cinque (2010 : 107), seul un adjectif articulé est toutefois « naturel ». Cinque (2010 : 107) parle de non itérabilité de l'adjectif articulé – comme en roumain. Les structures DEF N DEF A DEF A, DEF A DEF A DEF N, ou encore DEF A DEF N DEF A, seraient donc artificielles.

2.3. Analyse syntaxique

Il existe plusieurs analyses syntaxiques concurrentes du phénomène :

- Proposition relative réduite : Androutsopoulou (2001)
- Clitic doubling réalisant une tête fonctionnelle : Alexiadou (2014 : 29-30), à la suite d'Alexiadou et Anagnostopoulou (2000), Anagnostopoulou (2006), Anagnostopoulou (2013)
- Hypothèse de la scission du DP : Androutsopoulou (1995), Kariaeva (2004), Ioannidou et den Dikken (2009), Kyriakaki (2010), Kyriakaki (2011)
- Hypothèse de la prédication sous-jacente : Campos et Stavrou (2004)
- Ellipse du nom dans la sous-partie adjectivale : Panagiotidis et Marinis (2011), Lekakou et Szendrői (2012), Stavrou (1996)
- Apposition étroite : Lekakou et Szendrői (2007), Lekakou et Szendrői (2012), Lekakou et Szendrői (2014) Panagiotidis et Marinis (2011)
- FocP : Ntelitheos (2004), Kariaeva (2004), Alexiadou et Wilder (1998)

2.3.1. Accord en définitude

Le phénomène en définitude est traité par Corbett (2006 : 135-137). Selon Androutsopoulou (1995), la répétition de l'article défini consiste en la répétition du trait de définitude, qui est purement syntaxique et non sémantique. Les déterminants sont des instances de têtes fonctionnelles dans la projection Def° (définitude) du DP. Ce Def° est optionnellement projeté au-dessus du NP et de l'AP, hébergeant les traits de définitude et de cas. Les articles qui s'y trouvent sont donc explétifs et ne s'interprètent pas sur le plan sémantique. Ils ne sont pas sur le même plan que l'article « initial », qui n'est pas issu du *spreading*. Celui-ci se trouve dans la plus haute tête de la projection étendue du nom. S'il y a répétition de l'article en surface, il n'y a pas répétition de son contenu sémantique ; les instances répétées sont explétives ; il y a donc un statut différent et inégal entre l'article initial et les articles répétés.

Kariaeva (2004) suit cette théorie et soutient que le second article défini est un cas d'accord en définitude – accord absent des syntagmes monodéfinis. Il rappelle l'accord en définitude présent dans d'autres langues, telles les langues sémitiques (Rubin 2005b), dont l'hébreu (exemple (90).[1]

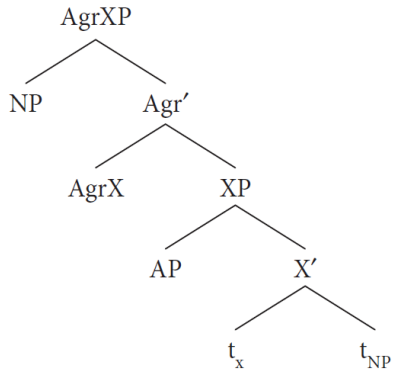
(90) Hébreu (Alexiadou 2014 : 80, 81, 83)

a. *ha dira ha gdola*

[1] On peut toutefois comparer le schéma de la structure sous-jacente en hébreu (exemple (90)c) avec celle supposée pour le grec en (69)b. L'arbre de Leu (2009) et celui d'Alexiadou (2014) diffèrent.

	the	apartment	the	big
b.	*ha*	*xatulim*	*ha*	*rabim*
	the	cats	the	many

c.



Lorsque la définitude est marquée plusieurs fois dans un même syntagme nominal sur des éléments de nature différente, tels que le nom et ses modifieurs, on peut parler d'accord en définitude. Androutsopoulou (1995 : 171-173) discute cette hypothèse, qui est très tentante en raison de l'identité morphologique de l'article « adjectival » avec l'article « nominal ». L'autrice propose de considérer qu'un article défini accompagne toujours l'adjectif épithète en grec, mais qu'il n'est pas réalisable du point de vue morphologique lorsque l'article « nominal » est indéfini (*IND N DEF A). Ce serait la seule occurrence, dans la langue, d'un tel accord… alors comment l'expliquer, et comment expliquer son optionnalité ? Cette théorie lui permet de supposer une structure sous-jacente commune au grec et à l'albanais – mais non en structure de surface (exemple (91). L'albanais permet en effet la présence du « déterminant » adjectival (obligatoire pour quasiment tous les adjectifs) dans les syntagmes indéfinis (exemple (92).

(91) Albanais (Alexiadou 2014 : 86)

një	*djalë*	*i*	*mirë*
a	boy	the	good

'a good boy'

(92) Grec (Alexiadou 2014 : 86)

**ena*	*agori*	*to*	*kalo*
a	boy	the	good

Lekakou et Szendrői (2014 : section 3.2) écrivent que le syntagme polydéfini du grec peut être un certain type d'accord en définitude, nommé *Definiteness concord generalisation*. « *In Def-D split languages, any nominal element in the scope of a definite operator must be marked for definiteness by the presence of the syntactic marker for definiteness, D* » (Lekakou et Szendrői 2014 : 224). Elles posent une règle valable pour les langues à scission Def-D : *Def [D NP … *(D) NP].

2.3.2. Proposition relative sous-jacente

L'autre hypothèse permettant d'expliquer la structure est de présumer qu'il s'agit d'une proposition relative dont tous les termes ne sont pas exprimés en surface. Alexiadou et Wilder (1998) proposent

d'analyser la partie « adjectivale » du syntagme comme une proposition relative sous-jacente, dont le sujet est un DP plutôt qu'un NP. Un ensemble de mouvements a lieu qui explique la forme de surface. Cette proposition relative serait de type prédicative, l'adjectif y occupant la fonction d'attribut du sujet et le nom y occupant la fonction de sujet : [DEF N] [DEF A] // [sujet] [prédicat]. Pour paraphraser avec un exemple français, on peut dire qu'un syntagme comme *le crayon jaune* ou **le crayon le jaune* aurait une structure sous-jacente *le crayon (qui est (le)) jaune*. La Figure 25 montre sa structure selon Campos et Stavrou (2004 : 157) :

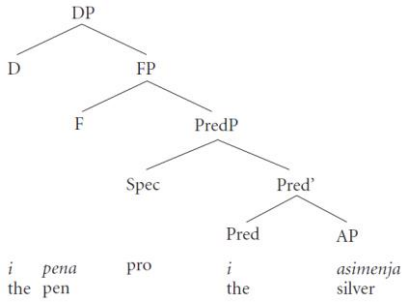


Figure 25 : Structure sous-jacente d'un syntagme polydéfini grec (Campos et Stavrou 2004 : 157)

Des critiques ont été émises sur la validité de cette structure sous-jacente, comme la génération du nom sous D, et la différence supposée entre la structure sous-jacente entre un syntagme monodéfini et un syntagme polydéfini (Kyriakaki 2011 : 51-53). Mais plusieurs auteurs, dont Alexiadou (2005 : 810), ont expliqué que la structure polydéfinie ne pouvait apparaître qu'avec des adjectifs épithètes pouvant être également prédicatifs. Dans les exemples suivants, l'adjectif ne peut être prédicatif, et donc ne peut entrer dans un syntagme polydéfini.

(93) Grec (Alexiadou et Wilder 1998 : 306)

	i	*italiki*	*(*i)*	*isvoli*		
	the	Italian	the	invasion		
*	*i*	*isvoli*	*stin*	*Alvania*	*itan*	*italiki*
	the	invasion	of	Albania	was	Italian

Par conséquent, les adjectifs qui ne peuvent revêtir la fonction d'attribut du sujet ne peuvent pas entrer dans un syntagme polydéfini (Alexiadou et Wilder 1998 : 306). Les adjectifs « non articulables » ne peuvent pas jouer la fonction d'attribut du sujet (Cinque 2010 : 106). L'inverse n'est pas vrai : il existe des adjectifs « articulables » qui ne peuvent pas non plus jouer la fonction d'attribut du sujet (Androutsopoulou 2001 : 191).

D'où les règles syntaxiques suivantes :

- « *All instances of (articled) adjectives in polydefinites can be used as predicates in a predicative clause (cf. also Alexiadou & Wilder's 1998 generalization (30)).* » (Campos et Stavrou 2004 : 156, Alexiadou 2014 : 32)
- « *The opposite does not hold, as there are predicative adjectives that cannot be freely employed preceded by an article, although they obviously have a predicative source. (cf. Manolessou 2000, Alexiadou 2007)* » (Alexiadou 2014 : 32)

Selon Alexiadou (2014 : 35-39, 52), si le *determiner spreading* provient d'une proposition relative réduite sous-jacente, il faut supposer une structure sous-jacente différente au syntagme monodéfini et au syntagme polydéfini. Celle-ci doit avoir une interprétation restrictive (Alexiadou 2014 : 35). « *It is predicted that adjectives that cannot be used predicatively (in copula sentences etc.) should not permit determiner spreading in attributive use* » (Alexiadou et Wilder 1998 : 306).

Cinque (2010 : 105-107) liste les caractéristiques syntaxiques et sémantiques de la structure polydéfinie grecque. Il en déduit que les adjectifs de telles structures prennent leur source dans une proposition relative réduite, mais en supposant que l'adjectif et son article constituent un DP avec ellipse du nom (théorie présentée en section 2.3.3, p. 166), DP qui serait le prédicat de la relative (Cinque 2010 : 107). À moins que le DP avec ellipse du nom ne soit une apposition étroite au DP contenant le nom modifié (théorie présentée en section 2.3.4.2, p. 173). Dans tous les cas, la partie adjectivale du syntagme polydéfini serait un cas de modification indirecte (Cinque 2010 : 108). Notons que l'auteur ne précise pas s'il existe une différence entre les structures polydéfinies à adjectif antéposé et à adjectif postposé – DEF A DEF N et DEF N DEF A seraient donc deux structures à modification indirecte. L'analyse du syntagme polydéfini grec comme un cas de modification indirecte est d'ailleurs cohérente avec son analyse du syntagme monodéfini. Lui et Stavrou notent de nombreuses similarités entre l'adjectif non articulé grec et l'adjectif anténominal anglais : tous deux sont ambigus quant aux deux types d'interprétation listés dans le Tableau 9. C'est donc plutôt au type germanique qu'au type roman qu'il faudrait assimiler le grec. La modification indirecte est celle qui est ambiguë quant à l'interprétation, et elle se trouve exclusivement en position nominale en anglais. En grec, elle se trouverait donc à la fois en syntagme monodéfini (position anténominale) et en syntagme polydéfini (position anté- et postnominale).

2.3.3. Ellipse du nom

L'absence de nom en présence d'un adjectif est possible en grec devant un article défini (94), un article indéfini (95) et l'article zéro (96). On peut considérer qu'il s'agit d'une substantivation de l'adjectif, qui devient alors la tête du syntagme nominal, ou bien qu'il s'agit d'une ellipse du nom. La seconde option est la plus plausible, car l'adjectif s'interprète en fonction du nom antécédent. Ainsi en (94), « la verte » ne désigne pas n'importe quel référent valable en tout contexte, mais une certaine robe propre au co-texte.

(94) Grec (Lekakou et Szendrői 2007 : 149)

I	*Maria*	*forese*	*to*	*ble*	*fustani*	*ke*	*i*	*Eleni*	*forese*	*to*	*prasino* *Ø.*
the	Maria	wore	the	blue	dress	and	the	Eleni	wore	the	green

[Ø = fustani]

(95) Grec (Lekakou et Szendrői 2012 : 123)

I	*Maria*	*agorase*	*ena*	*akrivo*	*forema*	*ke*	*i*	*Eleni*	*ena*	*ftino.*
the	Maria	bought	an	expensive	dress	and	the	Eleni	an	cheap

'Maria bought an expensive dress et Eleni a cheap one.'

(96) Grec (Lekakou et Szendrői 2012 : 124)

I	*Maria*	*agorazi*	*akriva*	*ruxa*	*ke*	*i*	*Eleni*	*ftina.*
the	Maria	buys	expensive	clothes	and	the	Eleni	cheap

'Maria buys expensive clothes and Eleni cheap ones.'

1) Panagiotidis (2005) et Panagiotidis et Marinis (2011) proposent une ellipse du nom dans la sous-partie « adjectivale » du syntagme polydéfini. Ils posent une hiérarchie entre les deux sous-parties, l'une étant sujet de l'autre dans une proposition relative réduite (RCC – *reduced relative clause*). Cette théorie résout le problème de l'ordre des mots et de la présence multiple de Det dans le syntagme polydéfini. Elle concernerait seulement l'ordre DEF N DEF A.

Cinque (2010 : 104-108) considère lui aussi l'adjectif articulé comme une proposition relative réduite, et que l'article et l'adjectif constituent un DP avec un nom elliptique constituant le prédicat de la relative. Pour paraphraser avec un exemple français, on peut dire qu'un syntagme comme *le crayon le jaune* aurait une structure sous-jacente *le crayon (qui est) le (crayon) jaune*. Il propose les arguments suivants :

- Les adjectifs articulés sont nécessairement restrictifs, sinon la structure est agrammaticale.
- Les adjectifs articulés sont interprétés de façon intersective.
- Les adjectifs articulés présentent un ordre libre, au contraire des adjectifs non articulés.
- Les adjectifs articulés ne peuvent pas être plus proches du nom que les adjectifs non articulés.
- Les adjectifs qui ne peuvent pas être articulés ne peuvent pas non plus être attributs du sujet. L'inverse est presque vrai – seuls quelques adjectifs peuvent être articulés sans pouvoir être pour autant employés comme attributs du sujet.
- Les adjectifs articulés impliquent un contraste sémantique.

Par ellipse du nom, on entend « *a Discourse Grammar phenomenon* […] *based on an anaphoric connection between the ellipsis-DP (i.e. the DP containing the ellipsis site) and the antecedent-DP* » (Cornilescu et Nicolae 2012 : 1070). Elle se produit toujours dans des syntagmes définis :

> « *Ellipsis presupposes the retrieval, in fact, the construction of a common topic, which is jointly denoted by the elided NP and by the antecedent: the common topic is the set of entities of NP-kind denoted by the antecedent* […] *in conjunction with the set of NP-entities denoted by the ellipsis-DP* […]. *Therefore, the antecedent-DP and the ellipsis-DP each denote different subsets of the common topic* […], *referred to by the identical NP.* […] *Nominal ellipsis (=NomE) thus presupposes the analysis of the two DPs into a first, common NP part, which denotes the common topic, and a second, dissimilar part, which shows which subsets are denoted by the antecedent and by the ellipsis-DP, respectively. Since it is the second DP which overtly shows its anaphoric dependence on the antecedent, it is the structure of the ellipsis-DP which makes ellipsis possible. More precisely, since the NP in the ellipsis-DP, like the NP in the antecedent-DP, denotes the common topic, the second occurrence of the topic-denoting NP must be definite to structurally express the anaphoric connection with the antecedent.* [...] *The comparison with the antecedent must identify not only the common NP, but also a contrastive constituent, which brings in new information and represents the subpart of the topic set expressed by the ellipsis-DP in contrast with the antecedent. The remnant is a semantic focus and, as first claimed by Merchant (2001), it undergoes syntactic focus movement, to a position above the determiner (in NomE).* » (Cornilescu et Nicolae 2012 : 1071)

2) Lekakou et Szendrői (2007), Lekakou et Szendrői (2012) et Lekakou et Szendrői (2014) fondent une théorie de la close apposition (section 2.3.4, p. 176) sur l'ellipse du nom. Pour paraphraser avec un exemple français, on peut dire qu'un syntagme comme *le crayon le jaune* aurait une structure sous-jacente *le crayon le (crayon) jaune*. La structure sous-jacente du syntagme polydéfini est celle d'un

grand syntagme nominal comprenant en réalité deux sous-syntagmes nominaux, avec deux déterminants et deux têtes nominales. La seule différence entre une apposition étroite nominale (97) et une apposition étroite adjectivale en syntagme polydéfini (98), est que la première est constituée de deux DP contenant tous deux un nom, tandis que le deuxième est constitué de deux DP, l'un contenant un nom, l'autre contenant un adjectif et un nom ellidé, phonétiquement nul (Lekakou et Szendrői 2007 : 130-136).

(97) Grec (Lekakou et Szendrői 2012 : 108)

(a)	*i*	*asimenia*	*i*	*pena*
	the.NOM	silver	the.NOM	pen
(b)	*i*	*pena*	*i*	*asimenia*
	the.NOM	pen	the.NOM	silver

'the silver pen'

(98)

(a)	*o*	*aetos*	*to*	*puli*
	the.MASC	eagle	the.NEUT	bird
(b)	*to*	*puli*	*o*	*aetos*
	the.NEUT	bird	the.MASC	eagle

'the eagle that is a bird'

L'ellipse permet d'expliquer l'effet sémantique que d'aucuns attribuent au focus (Lekakou et Karatsareas 2016 : 194) (cf. section 2.5, p. 192). Elle permet d'expliquer que l'adjectif semble articulé sans être suivi d'un nom, voire nominalisé. Elle explique aussi pourquoi l'ordre des mots est libre dans une construction polydéfinie : les deux sous-parties sont structurellement à égalité, sans hiérarchie entre elles. Elles sont sont totalement formées sur les plans sémantique et syntaxique (Lekakou et Szendrői 2007 : 139). La partie adjectivale est un DP *complet*, comprenant une ellipse du NP (Szendrői 2010 : 873). Les phénomènes d'accord dans les deux sous-parties confirment que la structure polydéfinie implique une structure symétrique. À cette restriction près : si un adjectif n'est pas précédé d'un article défini, il doit précéder le nom, en configuration monodéfinie classique.

(99) Grec (Lekakou et Szendrői 2007 : 143)

(a)	*	*to*	*spiti*	*megalo*	*to*	*petrino*
		the	house	big	the	stone
(b)	*	*to*	*spiti*	*to*	*megalo*	*petrino*
		the	house	the	big	stone

Si aucune partie ne constitue plus la tête que l'autre, une telle théorie est en contradiction avec celles exposées en section 2.3.2, p. 164. Lekakou et Szendrői (2012 : 133) montrent que l'analyse par proposition relative sous-jacente ne peut pas expliquer un certain nombre de faits :

- l'interprétation obligatoirement restrictive de l'adjectif,
- l'ensemble d'adjectifs et d'interprétations d'adjectifs possibles ou non dans la construction,
- les adjectifs pouvant entrer en construction polydéfinie dans certaines circonstances, mais qui ne peuvent pas se trouver en position prédicative,
- le problème de la *poly*-définitude : la réalisation multiple d'un déterminant défini qui n'est interprété qu'une seule fois.

3) Cornilescu et Nicolae (2012 : 1071) rejoignent les autrices en ce que l'ellipse présuppose une structure de DP identique au DP sans ellipse. Ils affirment toutefois que ce qui reste du DP après élision est focalisé contrastivement. Suivant Rizzi (1997), ils donnent des arguments structurels permettant de différencier un DP avec élipse du nom d'un DP sans élision. Kyriakaki (2010 : 10-11) montre que la tête du modifieur polydéfini est un nom vide : elle rejoint les partisans de l'ellipse, mais avec plusieurs bémols. Aucun élément ne peut se glisser dans la partie adjectivale ; rien ne peut venir s'intercaler entre l'adjectif et le nom, ni adjectif, ni adverbe, ni syntagme génitival. La partie adjectivale est structurellement très rigide : elle forme un DP « petit » ou « réduit ». Seul peut s'intercaler un pronom personnel clitique[1] permettant d'exprimer la possession, comme dans l'exemple (100).

(100) Grec (Kyriakaki 2010 : 10)

to	*kokino*	*tis*	/ *tis	Marias	*to*	*poðilato*
the-NEU	red-NEU	her-CLIT	the-GEN.FEM	Maria-GEN.FEM	the-NEU	bike-NEU

'Her red bike'/ *Mary's red bike (ungrammatical for Greek)'

2.3.4. Apposition

L'apposition est un syntagme qui en modifie un autre, dont il partage le même référent, auquel il apporte une précision, une information supplémentaire. On parle également d'épithète détachée du nom – mais non détachée du reste du syntagme nominal (Haßler 2016 : 28). Elle peut être placée avant ou après le syntagme modifié, nommé syntagme support. Plusieurs auteurs ont analysé la structure grecque comme un cas d'apposition, notamment Lekakou et Szendrői (2007), Lekakou et Szendrői (2009), Lekakou et Szendrői (2012) et Lekakou et Szendrői (2014).

2.3.4.1. Deux types d'apposition

Il convient de différencier tout d'abord les appositions étroites (*close*, exemple (101)a) et lâches (*loose*, exemple (101)b) :

(101) Anglais (Lekakou et Szendrői 2007 : 131)

a. *Burns the poet* (close apposition)

b. *Burns, the poet* (loose apposition)

La différence entre apposition étroite et lâche est le type de lien qui les unit, donc l'autonomie plus ou moins importante dont dispose le syntagme apposé.

Une **apposition étroite** est directement adjacente au syntagme qu'elle modifie, sans pause. Elle ne comporte pas de rupture d'intonation et ne peut porter qu'un seul accent. À l'écrit, l'ensemble est donc écrit sans séparation. Aucun mot ne peut s'intercaler entre l'apposition et le syntagme qu'elle

[1] Il n'existe pas de déterminant possessif en grec : on associe l'article défini et le pronom personnel clitique.

modifie. L'apposition étroite n'est possible que dans le domaine nominal, avec des éléments de même nature (Alexiadou 2014 : 46). Les deux forment un grand syntagme nominal commun, une seule unité prosodique. Dans une apposition étroite, les deux parties nominales ont un statut égal et contribuent à la détermination et à l'identification de la référence : les deux DP sont donc référentiels (Lekakou et Szendrői 2012 : 111). Selon Lekakou et Szendrői (2012 : 139), la preuve que les deux DP forment un unique constituant DP est qu'ils reçoivent tous deux le morphème de cas et la préposition, voire le déterminant. Ils sont symétriques et complets sur les plans sémantique et syntaxique.

Une apposition étroite doit être définie ; ses deux sous-parties doivent contenir un déterminant défini. Dans la structure polydéfinie, la sous-partie « adjectivale » contient un article, un adjectif et un nom ellidé (Lekakou et Szendrői 2007), ou bien un article et un adjectif qui forment le prédicat d'une proposition relative réduite (Cinque 2010 : 107). Elle modifie le nom de la sous-partie « nominale » (Panagiotidis et Marinis 2011, Cinque 2010 : 107).

Les deux sous-parties d'une apposition étroite ne peuvent être que deux syntagmes nominaux (plus précisément, deux syntagmes déterminatifs, DP). On trouve de telles appositions avec les noms propres (qui sont articulés en grec), ainsi qu'avec les titres, les termes d'adresse, les noms déterminatifs ou classifieurs, ou encore les surnoms (Plank 2003b : 342-343, exemple (102)a, b, c).

(102) Grec (Plank 2003b : 341-342)

a.	*o*	*θíos*	*/ fílos*	*mu*	*(o)*	*aléksis*		
	the	uncle	/ friend	of.me	the	Alexis		
	« (my) uncle Alexis / my friend Alexis »							
b.	*i*	*póli*	*i*	*Aθína*				
	the	town	the	Athens				
c.	*o*	*kírios*	*(*o)*	*zórbas;*	*o*	*aléksis*	*(*o)*	*zórbas*
	the	mister	(*the)	Zorbas;	the	Alexis	(*the)	Zorbas
	« Mr Zorbas ; Alexis Zorbas »							

L'article défini peut même être doublé lorsqu'il existe une ambiguïté sur l'identification de la personne dénotée par le nom (Chatzikyriakidis 2016 : 5-6, Kyriakaki 2011 : 18-19) :

(103) Grec (Chatzikyriakidis 2016 : 5)

I	*Maria*	*i*	*Papadopulu*	*ine*	*fili*	*mou*
the	Maria	the	Papadopoulou	is	friend	my

'Maria Papadopoulou is a friend of mine.'

En fonction du contexte, le syntagme sert à distinguer un individu spécifique, Maria Papadopoulou, au sein de l'ensemble des personnes portant le nom de famille Papadopoulou ou bien des personnes portant le prénom Maria (Chatzikyriakidis 2016 : 5-6).

Pour une présentation synthétique de l'apposition étroite en grec, ainsi qu'une analyse syntaxique, on pourra consulter Marinis (2003 : sections 7.1 et 7.2) et Alexiadou (2014 : 42-47). Cette dernière traite les cas d'apposition nominale étroite de sens péjoratif (dont le « deuxième » article défini est qualifié de *spurious article*, article fallacieux), les constructions partitives et pseudo-partitives (voir aussi Alexiadou 2014 : 50-51). Il existe plusieurs termes équivalents pour l'apposition étroite (Lekakou et Szendrői 2007). Plusieurs sont également employés pour désigner les deux parties de l'apposition lâche, qui sont hiérarchisées, contrairement à celles de l'apposition étroite. On notera l'ambiguïté du terme « apposition », qui désigne soit l'ensemble du syntagme appositionnel DEF N DEF N (ou DEF N DEF A), étroit ou lâche, mais aussi la sous-partie secondaire propre à l'apposition

lâche. Le Tableau 16 présente une synthèse des termes utilisés dans les travaux scientifiques.

Tableau 16 : Les différents noms donnés aux types d'apposition et à leurs sous-parties

Apposition étroite	**Apposition lâche**	
close apposition	*loose apposition*	
restrictive apposition	*non-restrictive apposition*	
integrated appositive	*supplementary appositive*	
	partie 'principale'	**partie 'secondaire'**
	host	*appositive*
	anchor	*apposition*
		supplement

Je choisis de traduire *close apposition*, « apposition proche/fermée », par apposition **étroite**, *loose apposition*, « apposition lâche/large », par apposition **lâche**, *anchor* par **ancre** et *appositive/apposition* par **partie appositive** – afin d'éviter l'ambiguïté.

L'**apposition lâche** est un phénomène consistant, pour un syntagme appositif, à en modifier un autre (l'ancre) de manière adjacente ou non. Le syntagme principal est nommé ancre et le syntagme secondaire est nommé partie appositive. Il s'agit de deux unités juxtaposées, qui peuvent porter chacune leur propre accent. La partie appositive forme un syntagme indépendant de l'ancre. L'apposition lâche comprend une rupture ou un changement intonatif entre les deux parties du syntagme. Lekakou et Szendrői (2007) et Lekakou et Szendrői (2012 : 111) parlent respectivement d'*intonational pause* et d'*intonational break.* Les deux syntagmes forment donc deux unités prosodiques différentes. L'apposition lâche peut comprendre également une longue pause, transcrite à l'écrit par la ponctuation qui l'encadre : virgules, tirets, parenthèses, deux points (Alexiadou 2014 : 34). Elle se rencontre avec l'article défini (104) et l'article indéfini (105).

(104) Grec (Alexiadou 2014 : 34)

To	*spiti*	*to*	*kokkino,*	*to*	*megalo,*	*itan*	*poli*	*akrivo.*
the	house	the	red,	the	big,	was	very	expensive

'The red house, the big one, was very expensive.'

(105) Grec (Alexiadou 2014 : 34)

Fere	*mou*	*ena*	*vivlio,*	*ena*	*kalo!*
Bring	me	a	book	a	good

'Bring me a book, a good one!'

Elle peut comprendre enfin l'insertion d'éléments lexicaux à la frontière entre les deux sous-parties : la partie appositive est séparée de l'ancre par un adverbe modalisateur, une interjection, un terme métalinguistique, une locution, un ou des syntagmes, voire tout un énoncé. C'est la présence de la frontière prosodique qui rend possible cette insertion selon Lekakou et Szendrői (2007 : 131).

(106) Anglais (Lekakou et Szendrői 2007 : 131)

a.		*the head of department, namely Prof. Todorov*	apposition lâche
b.	*	*Burns namely the poet*	apposition étroite

L'apposition lâche est possible avec n'importe quelle catégorie grammaticale, y compris avec des éléments de nature différente (Alexiadou 2014 : 46) dans les deux sous-parties : verbe + verbe, syntagme prépositionnel + syntagme prépositionnel, adjectif + adjectif, nom + nom, proposition + proposition (Lekakou et Szendrői 2007 : 132). Elles n'ont pas de lien syntaxique étroit et peuvent être séparées l'une de l'autre. La partie appositive fonctionne comme un ajout ou une parenthèse. L'apposition lâche est d'ailleurs considérée comme une structure parenthétique par (Lekakou et Szendrői 2012 : 111).

L'ancre est dite référentielle, la partie appositive est dite prédicative. Contrairement à l'apposition étroite, l'information de l'apposition lâche est supprimable sans altération de la grammaticalité de la phrase. C'est qu'elle ne contribue pas à la valeur référentielle de l'ancre, le syntagme support. Enfin, la référence ne fonctionne pas de la même manière dans les deux types d'apposition (Lekakou et Szendrői 2007 : 132-133). Dans l'apposition lâche, l'ancre sélectionne une entité unique et la partie appositive fournit une information supplémentaire sur celle-ci (exemple (107). L'ancre revêt donc une plus grande importance et respecte le principe d'unicité. Dans l'apposition étroite, les deux sous-parties de l'apposition contribuent ensemble à la construction de la référence. L'interprétation n'est donc pas restrictive dans l'apposition lâche, c'est-à-dire que l'adjectif ne sert pas à identifier le référent parmi un ensemble d'autres référents possibles :

(107)		Anglais (Kyriakaki 2010 : 4)	
	a.	My sister Elena lies in Berlin and my sister Eleftheria in Athens.	apposition étroite
	b.	My sister, Eleftheria, lives in Athens.	apposition lâche
(108)		Anglais (Lekakou et Szendrői 2007 : 144)	
		Tonight I will speak of a great French artist.	
	a.	Guillem, the dancer … .	apposition étroite
	b.	# Guillem the dancer … .	apposition lâche

L'exemple (108)a est grammatical car il fait référence à une entité unique du monde, désignée par la première partie de l'apposition, *Guillem*. Il n'implique pas de restriction sémantique. Au contraire, l'exemple (108)b n'est pas grammatical dans ce contexte car la première partie de l'apposition, *Guillem,* ne présuppose pas que Guillem soit une entité unique du monde. L'exemple serait grammatical s'il existait plusieurs Guillem, d'où le besoin de préciser que l'on parle de celui qui est danseur. La référence s'établit grâce aux deux sous-parties de manière conjointe, et non par la première partie seulement. Les deux sous-parties, prises séparément, violent le principe d'unicité ; celui-ci est respecté seulement au niveau de l'ensemble de la construction.

Enfin, notons que dans certaines langues, la présence d'une virgule ne signifie pas forcément que l'apposition soit lâche. Ainsi existe-t-il en français deux types de virgules correspondant à deux effets structuraux différents (Nølke 2001 : 211) : « les **virgules-c**, qui marquent une coordination et les **virgules-r** qui marquent une rupture intonative, et par là une frontière syntaxique majeure ».

- Les virgules-c sont une simple juxtaposition de modifieurs en apposition étroite, ou épithètes appartenant à la « post-zone » (notion définie en section 3.5, p. 97) : il s'agit d'une **coordination asyndétique** (Nølke 2001 : 216).

- Les virgules-r sont des **détachements** en apposition lâche, prononcés « avec l'intonème parenthétique » (Nølke 2001 : 212). Elles marquent un changement de fonction syntaxique.
- Il existe d'autres emplois de la virgule, qui transcrit un « emploi particulier de l'intonation : le deuxième adjectif est ajouté comme une sorte d'après-coup », ainsi dans l'exemple « *une atmosphère trop sévère, trop médiévale pour le commun des mortels* », ou encore *son bruit ferrugineux, intarissable et glacé* (Nølke 2001 : 218, 222).

2.3.4.2. Apposition étroite

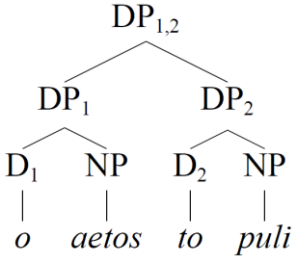
Les appositions étroites sont restrictives, au contraire des appositions lâches qui impliquent une présupposition d'unicité. Les structures polydéfinies se comportent comme les premières et non comme les secondes.

Stavrou (1991), puis (Kolliakou (1998 : 19-20) et (Kolliakou (2004) décident de mettre sur le même plan les adjectifs des syntagmes polydéfinis (DEF N DEF A et DEF A DEF N) et, entre autres, les appositions nominales étroites qui sont classiquement nominales (DEF N DEF N et DEF N DEF N) :

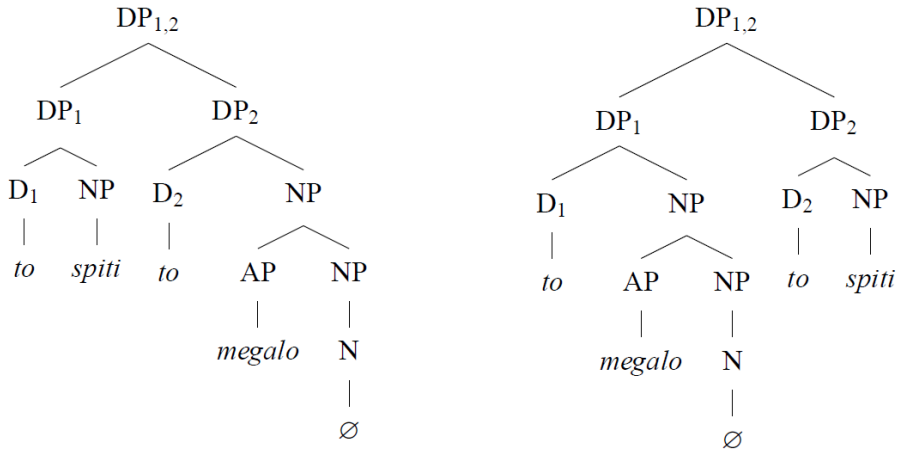
> « *An analysis a la Sag 1997 in terms of construction type specification, allows us to generalize over the commonalities between definite APs of polydefinites and other types of 'adnominal' modifiers, e.g. (a) predicative adjectives in the sense of Bolinger, as in* the rivers, impassable in the winter, … *(noticeably, predicative APs in Greek can appear both pre- and post-nominally, exactly like definite APs in polydefinites, and it has been previously proposed that they should be viewed as modifiers of maximal nominal projections, cf. Stavrou 1996); (b) NPs in close apposition, e.g. my brother the lawyer, and (c) participles functioning as reduced relatives, e.g. a man sitting on the bench. NP-internal adjectives of monadics (which have been argued independently to be syntactic heads rather than adjuncts, cf. Kolliakou 1998) are constrained (by a linear precedence statement) to preceed their nominal (NP or AP) sister; on the other hand, adnominal modifiers of polydefinites are unconstrained and can preceed or follow the NP head.* » (Kolliakou 1998 : 19-20)

Comparons la structure sous-jacente d'une apposition étroite DEF N DEF N, présentée en (109), et la structure sous-jacente d'un syntagme polydéfini DEF N DEF A ou DEF A DEF N, présenté en (110), telle que présentée par Lekakou et Szendrői (2007). Les autrices révisent les arbres dans un article postérieur (111) et (112). La structure est la même pour l'apposition étroite nominale et pour le syntagme polydéfini. Aucune des deux sous-parties n'étant la tête de la construction générale, il n'existe pas d'ordre intrinsèque entre elles, ce qui explique l'ordre libre DEF N DEF A ou DEF A DEF N dans les syntagmes polydéfinis.

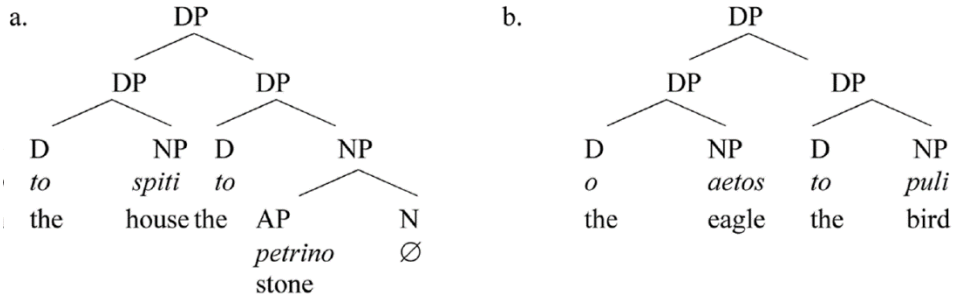
(109) Grec (Lekakou et Szendrői 2007 : 135)



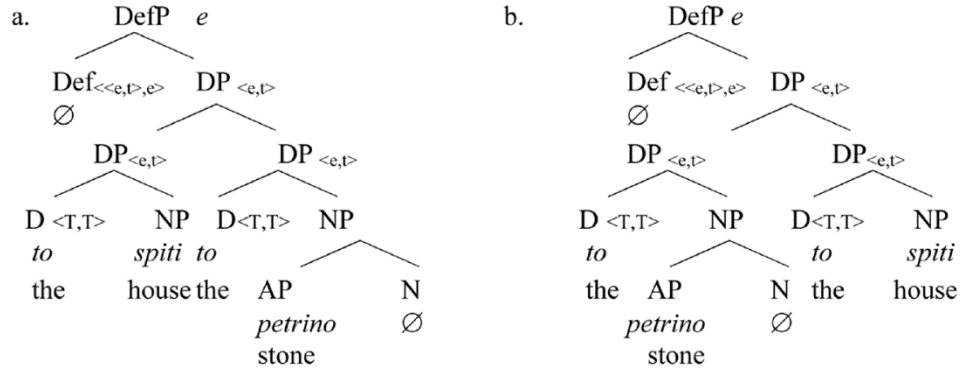
(110) Grec (Lekakou et Szendrői 2007 : 140)



(111) Grec (Lekakou et Szendrői 2014 : 218)



(112) Grec (Lekakou et Szendrői 2014 : 220)



Lekakou et Karatsareas (2016 : 194), travaillant sur deux variétés différentes de grec, montrent que l'apposition étroite nominale présente les mêmes caractéristiques syntaxiques et sémantiques que le syntagme polydéfini adjectival, et plaident par conséquent pour une définition globale de l'apposition étroite, adjectivale ou nominale, à la structure similaire.

En effet, le syntagme polydéfini et l'apposition étroite partagent de nombreux points communs. L'ordre des mots, plus exactement, l'ordre des deux sous-parties, est libre. Les deux sous-parties sont introduites par un article défini. Dans les deux cas, un article indéfini est agrammatical (exemples (113), (114) et (115)), voir Lekakou et Szendrői (2012 : 110) et Alexiadou (2014 : 46). La polyindéfinitude, ou encore marquage multiple de l'indéfinitude, est donc impossible parce qu'il relève d'une loi plus générale, celle de l'apposition étroite.

(113) Grec (Lekakou et Szendrői 2012 : 123)

a.	*ena*	*megalo*	*(*ena)*	*spiti*
	a	big	a	house
b.	*ena*	*spiti*	*(*ena)*	*megalo*
	a	house	a	big

(114) Grec (Lekakou et Szendrői 2012 : 124)

a.	*	*o*	*Nikos*	*enas*	*kathijitis*
		the	Nikos	a	professor
b.	*	*enas*	*kathijitis*	*o*	*Nikos*
		a	professor	the	Nikos

(115) Grec (Lekakou et Szendrői 2012 : 110)

a.	*	*enas*	*aetos*	*(ena)*	*puli*
		an	eagle	a	bird
b.	*	*ena*	*puli*	*(enas)*	*aetos*
		a	bird	an	eagle

On constate que l'apposition étroite nominale (DEF N DEF N, exemple (98)) et l'apposition étroite adjectivale (DEF N DEF A ET DEF A DEF N) partagent la même morphologie (marquage du cas, double marquage de l'indéfinitude impossible) et la même syntaxe (ordre des mots plus libre que dans les cas de marquage simple de la définitude). Les deux partagent la même propriété de récursivité (Alexiadou 2014 : 46). Elles partagent aussi la même sémantique, celle de l'interprétation restrictive : un élément restreint l'interprétation de l'autre (Alexiadou 2014 : 45-46). Le modifieur restrictif, la partie adjectivale du syntagme polydéfini, est en effet essentielle sur le plan sémantique pour le sens de la phrase (Kyriakaki 2011 : 95), ainsi dans l'exemple (116).

(116) Grec (Kolliakou 2004 : 270)

a. *O diefthindis dilose oti <u>i kali erevnites</u> tha eprepe na apolithun.*
'The director declared that the competent researchers should be fired. ' (Two readings)

b. *O diefthindis dilose oti <u>i kali i erevnites</u> tha eprepe na apolithun.*
'The director declared that the competent researchers should be fired. ' (Restrictive reading only)

Elle est en revanche optionnelle sur le plan syntaxique, à l'instar des adjectifs « ordinaires » (Kyriakaki 2011 : 100). Cela explique pourquoi la structure est récursive. L'apposition restrictive sert à identifier le nom en restreignant sa dénotation, en réduisant l'extension de l'ensemble de ses référents possibles. Elle offre une information nouvelle, qui restreint l'ensemble (Kyriakaki 2011 : 101). Les modifieurs adjectivaux qui ne sont pas des DP mais des adjectifs nus diffèrent cependant dans leur structure sous-jacente (Kyriakaki 2011 : 101). Enfin, cette théorie présente l'avantage de supposer une structure sous-jacente de même type en syntagme monodéfini et polydéfini.

Kyriakaki (2011 : 56-57) émet toutefois des critiques face à cette approche, montrant qu'un DP restrictif et une proposition relative modifient le nom de la même manière, ce qui devrait impliquer une structure similaire. Kyriakaki (2011 : 57) critique également le fait que les deux DP soient des branches sœurs. L'une des deux branches est supérieure à l'autre : il s'agit de la « matrice », la partie contenant le nom en surface, car c'est avec celle-ci que s'accorde un éventuel attribut dans la phrase. La structure interne du syntagme polydéfini est asymétrique et la partie « nominale » en est la tête.

Alexiadou (2014 : 51) émet elle aussi des critiques contre la théorie de l'apposition étroite, qui prédit que le *determiner spreading* devrait être possible avec tous les adjectifs autorisant une ellipse. Ce n'est pas le cas des quantifieurs et des numéraux, qui autorisent l'ellipse mais ne peuvent faire

l'objet du determiner spreading. Cette théorie prédit également, selon elle, que le *determiner spreading* devrait être possible avec tous les types d'adjectifs prédicatifs, ce qui n'est pas le cas. La théorie de la proposition relative sous-jacente ne présente pas, selon elle, ce problème. Cette dernière permet au contraire d'expliquer l'asymétrie entre les deux sous-parties du syntagme polydéfini du point de sur le plan sémantique (Alexiadou 2014 : 52), ainsi que la possibilité de *determiner spreading* avec d'autres éléments que l'article défini : le numéral « un » et l'article indéfini Ø (Alexiadou 2014 : 52, Alexiadou 2014 : section 5.5). Elle propose au contraire de traiter les appositions étroites DEF N DEF N « classiques » (exemple (98) comme des propositions relatives sous-jacentes (Matushansky (2008) suit le même raisonnement).

2.3.4.3. Apposition lâche

Dans le cadre de la Syntaxe Dynamique, (Chatzikyriakidis 2014, Chatzikyriakidis 2015) propose l'existence de syntagmes polydéfinis ayant la structure d'une apposition lâche, c'est-à-dire pouvant présenter une pause ou une insertion lexicale entre les deux sous-parties. Il parle de *loose polydefinites*, polydéfinis lâches. Il montre que, si les syntagmes polydéfinis (exemple (118) sont des instances de *close appositions*, à l'instar de certains syntagmes nominaux apposés (117), on peut parler pour certains de *loose polydefinites*, polydéfinis lâches (120), qui sont des instances de *loose appositions*, appositions lâches (119).

(117) Grec (Chatzikyriakidis 2015 : 7)

a.	*o*	*aetos*	*to*	*puli*
	the	eagle	the	bird
b.	*to*	*puli*	*o*	*aetos*
	the	bird	the	eagle

(118) Grec (Chatzikyriakidis 2015 : 1)

a.	*to*	*spiti*	*to*	*megalo*
	the	house	the	big
b.	*to*	*megalo*	*to*	*spiti*
	the	big	the	house

(119) Grec (Lekakou et Szendrői 2007 : 136)

a.	*i*	*sikaminja*	*i*	*murja*
	the	blueberry tree	the	mulberry tree
b.	*i*	*murja*	*i*	*sikaminja*
	the	mulberry tree	the	blueberry tree

(120) Grec (Chatzikyriakidis 2015 : 12)

a.	*o*	*filos*	*mu,*	*o*	*omorfos*	
	the	friend	my,	the	handsome	
b.	*?o*	*omorfos,*	*o*		*filos*	*mu*
	the	handsome,	the		friend	my

Il existe donc une symétrie entre les quatre types de structures, apposition étroite et lâche, syntagme nominal ou nominal-adjectival avec polydéfinitude. Les syntagmes polydéfinis en apposition lâche remplissent la case manquante du Tableau 17.

Tableau 17 : Synthèse des différents types d'appositions en grec moderne

Appositions étroites		**Appositions lâches**	
Apposition nominale	Syntagme polydéfini	Apposition nominale	Syntagme polydéfini
(117)a. DEF N_1 DEF N_2 (117)b. DEF N_2 DEF N_1	(118)a. DEF N DEF A (118)b. DEF A DEF N	(119)a. DEF N_1, DEF N_2 (119)b. DEF N_2, DEF N_1	(120)a. DEF N, DEF A (120)b. DEF A, DEF N

L'ordre des constituants est libre, la partie adjectivale pouvant être antéposée (122) ou postposée (123) au nom. L'auteur, qui produit ses exemples, laisse toutefois planer un doute sur la grammaticalité de (120)b, avec adjectif antéposé. L'ordre des mots serait moins souple dans le syntagme polydéfini lâche que dans le syntagme polydéfini étroit, comme en témoigne également l'exemple (121)b.

(121) Grec (Chatzikyriakidis 2015 : 12)

a. *i Elines, i aneprokopi ine ipefthini*
the Greeks the shiftless are responsible
'The shiftless Greeks are to blame'

b. *?i aneprokopi, i Elines ine ipefthini*
the shiftless the Greeks are responsible
'The shiftless Greeks are to blame'

Comme dans une apposition lâche « classique », nominale, les deux sous-parties sont séparées à l'écrit par une virgule, à l'oral par une pause, un changement intonatif et/ou une insertion. L'interprétation des syntagmes polydéfinis en (122) et (123) n'est pas restrictive : l'ensemble des Chinois sont considérés comme inventifs et maléfiques, l'ensemble des Grecs considérés comme paresseux et bons à rien.

(122) Grec (Chatzikyriakidis 2015 : 2)
Pios itan aftos o laos? I efevritiki, i demonii, I Kinezi.
Who was this the people the inventive the Fiendish the Chinese
'Who were these people? The inventive and fiendish Chinese.'

(123) Grec (Chatzikyriakidis 2015 : 3)
I Elines, i tempelides, i aneprokopi ine ipefthini.
the Greeks the lazy the shiftless are responsible
'The lazy and shiftless Greeks are to blame.'

Un modifieur non-restrictif est selon Kyriakaki (2011 : 104) un ajout (*adjunct*) au nom. Il est optionnel tant sur le plan sémantique que sur le plan syntaxique, c'est pourquoi d'autres auteurs parlent également de « commentaire » ou de « parenthèse ». Il peut régir son propre modifieur, et on ne le trouve qu'en position postnominale.

2.4. Analyse sémantique

Pour Androutsopoulou (1995) ou Alexiadou et Wilder (1998), il n'y a pas de profonde différence de sens entre syntagme monodéfini et polydéfini. Ce point de vue est contredit par d'autres auteurs comme Kyriakaki (2011 : 48), selon laquelle les constructions polydéfinies ont une raison d'être : celle de transmettre un certain message et de révéler certaines fonctions du déterminant défini.

2.4.1. … de l'article : la surdétermination

Si l'ajout d'un possessif ou d'un démonstratif à un syntagme nominal défini lui confère un sens possessif ou démonstratif supplémentaire, l'ajout d'un article défini à un syntagme déjà défini devrait lui conférer un sens défini supplémentaire. C'est en ce sens que l'on pourrait parler de surdétermination, c'est-à-dire de détermination multiple avec une notion d'excès, ici de définitude plus qu'il n'est nécessaire.

Selon Androutsopoulou (1995), la répétition de l'article défini consiste en la répétition du trait de définitude, qui n'est pas sémantique. Le trait [+def], qui se manifeste par l'article défini, et qui est répété, n'est pas un trait porteur de sens. Sa répétition n'amène donc pas d'ajout de sens, pas de sur-détermination au sens propre : un tel syntagme n'est pas plus défini qu'un autre. En revanche, elle n'exclut pas que la motivation qui provoque l'emploi de cette structure soit d'ordre sémantique.

Plank (2003b : 383-386) dresse une étiologie de l'excès et une liste de sept bonnes raisons de « surmarquer » un syntagme :

- Lorsqu'un constituant est constitué de plusieurs parties qui sont de la même nature que l'ensemble, il peut être pertinent d'assigner les mêmes marques aux parties qu'à l'ensemble. C'est ainsi que le cas peut être répété. La distinction entre un NP lâchement relié à un autre, comme une apposition ou un *afterthouht*, et un NP emboîté dans un autre pour former un grand NP, bref la distinction entre apposition lâche et apposition étroite, n'est pas une distinction catégorique selon (Plank 2003b : 383). L'articulation serait sensible à l'intégrité du NP et serait prête à se répéter chaque fois qu'un adjectif épithète gagne en statut nominal ou commence à se détacher de la tête, ou change de sa position habituelle.
- Certaines catégories peuvent être utilisées pour indiquer la relation syntagmatique entre des parties de propositions. On utilisera volontiers des affixes pour relier l'accord entre elles, ou pour montrer qu'une partie commande une autre.
- Une catégorie peut être marquée plusieurs fois sur le même constituant pour exprimer un sens de façon compositionnelle, résultant de la combinaison des deux marqueurs. C'est le cas de la répétition du marqueur de pluriel, qui peut résulter en un paucal ou un pluriel, ou d'une marque de singulier combinée à une marque de pluriel, qui peut résulter en un sens de spécificité dans certaines langues.
- Un double marquage peut n'être qu'apparent, si les deux marqueurs n'ont pas la même fonction et qu'ils sont homonymes. Lorsqu'un nom porte deux marques de cas, l'un d'entre

eux peut en fait être un affixe ou porter un sens de topicalité. C'est le cas du superlatif en français, qui consiste à répéter ce qui semble la même chose : l'article défini, à ceci près que celui qui accompagne l'adjectif n'a pas la même fonction que celui qui accompagne le nom (*le livre le plus intéressant*). L'identité est formelle mais non fonctionnelle.

- La répétition d'un marqueur, comme celle d'un mot ou d'un groupe de mots, peut avoir une raison pragmatique plus que sémantique : emphase sur la catégorie concernée, force affective sur l'ensemble du syntagme…
- Un constituant peut être doublement marqué en raison de la réalisation simultanée de deux marquages équivalents mais syntagmatiquement exclusifs. Deux constructions alternatives, avec les mêmes éléments dans un ordre linéaire différent ou avec des éléments différents dans la même position, peuvent fusionner ou se contaminer. Ainsi le double marquage du pluriel peut-il provenir de la combinaison séquentielle de deux allomorphes du pluriel.
- Lorsqu'un marqueur cesse d'être distinctif, d'être compris ou reconnu, en raison de son érosion phonologique ou sémantique, ou bien en raison de sa réanalyse en autre chose, le constituant va recevoir une nouvelle marque de la même catégorie, pour des raisons de transparence. Ainsi, lorsqu'un article ou une partie d'un article a été amalgamé avec le nom et qu'il ne peut plus être analysé ni compris en synchronie, une réarticulation a lieu : la double articulation du nom n'en est plus une en synchronie, mais seulement en diachronie.

Plank (2003b) considère cependant la double articulation d'un nom comme quelque chose d'extra-ordinaire, qui ne peut se produire sans une bonne raison. Un SN doublement articulé cache souvent un autre SN en son sein, ou quelque chose qui est prêt à devenir un SN, ou qui jouit d'une certaine autonomie. Seule une telle raison syntaxique peut justifier l'effort extra-articulatoire. L'autre raison est d'ordre pragmatique : l'affectivité.

Selon Lebas-Fraczak (2015 : 45), la répétition non nécessaire d'un élément « qui aurait pu faire l'objet d'une ellipse » revêt une signification particulière du point de vue communicatif, au même titre que l'ellipse elle-même ou que « le choix de ne pas exprimer un élément ». La répétition exprime l'intensité, en lien avec la focalisation. En effet, plus un élément linguistique est étendu, « plus on le dote de "poids communicatif", ou de focalisation, au sein de l'énoncé » (Lebas-Fraczak 2015 : 46).

2.4.2. … de l'adjectif : l'intersectivité et la restrictivité

La plupart des auteurs s'accordent sur le fait que les adjectifs des syntagmes polydéfinis doivent entraîner une interprétation intersective du syntagme nominal (Kolliakou 1998 : 18, Kolliakou 2004 : 264, 267, Campos et Stavrou 2004 : 144, Cinque 2010 : 105).

(124) Grec (Campos et Stavrou 2004 : 144)

*	*o*	*monos*	*tu*	*o*	*erotas*	*ine*	*i*	*dulja*	*tu.*
	the	only	his	the	love	is	the	work	his

'His only love is his work.'

L'adjectif *monos*, seul, appartient selon les auteurs à la classe des adjectifs intensionnels, qui ne se prête pas à l'intersectivité. Ils ne permettent pas de sélectionner un référent qui fasse partie d'un ensemble d' *{entités seules} et de l'ensemble des {amours}. L'interprétation doit être conjointe, selon les termes de (Stavrou 1999 : 203) (Figure 13).

Beaucoup d'analyses montrent également que le type d'adjectif pouvant entrer dans une structure polydéfinie ne peut être que restrictif (Campos et Stavrou 2004, Lekakou et Szendrői 2007, Lekakou et Szendrői 2012, Cinque 2010 : 105, Kyriakaki 2011). Mais *ipotithemeni* ne peut pas dénoter l'ensemble des terroristes, ni dénoter les entitées prétendues ; il peut seulement identifier des entités prétendues parmi les terroristes. Qu'entend-on par *restrictif* ? (Givón 1990 : 10) donne la définition suivante :

> « *Restrictive modifiers are used to narrow down the domain of reference, in this way resembling definite determiners. The head noun may be semantically compatible with other referents in the relevant discourse context, and this may interfere with unique identification. The restrictive modifier rules out such conflicting interpretations.* […] *Non-restrictive modifiers are used to enrich the description of the referent with further characteristic attributes, but without narrowing down the domain of reference*. » (Givón 1990 : 10)

Une interprétation est dite restrictive lorsqu'elle identifie le référent parmi un ensemble d'autres référents possibles. L'adjectif vient restreindre l'extension du nom (Alexiadou 2005 : 810). La sémantique particulière des adjectifs intersectifs suppose une restriction pragmatique parmi les référents possibles. L'interprétation du syntagme polydéfini est restrictive, c'est-à-dire que l'on identifie un seul référent parmi un ensemble d'autres référents possibles. Une telle restriction est une relation entre la sémantique de l'adjectif et celle du nom. C'est l'adjectif qui vient ainsi restreindre l'extension du nom. La fonction du syntagme polydéfini est de « signifier une restriction nominale au nom-tête » (Kyriakaki 2011 : 134). La lecture du syntagme est hiérarchique, c'est pourquoi son interprétation est restrictive. Les adjectifs permettant une restriction pragmatique confèrent le plus souvent une lecture intersective (ou en intersective en apparence subsective) au syntagme nominal (Lekakou et Szendrői 2007 : 147).

Une interprétation est dite non-restrictive lorsqu'elle ne sert pas à identifier le référent parmi un ensemble d'autres référents possibles. L'adjectif y dénote une propriété qui prend une autre propriété en tant qu'argument, utilisée pour dénoter un type d'entité ou d'événement : c'est notamment le cas des adjectifs ni intersectifs ni subsectifs. Voyons des exemples (125) et (126) explicités par Cinque (2010 : 8) et Cinque (2014 : 4).

(125) Italien (Cinque 2010 : 8), (Cinque 2014 : 4)

a.	*Le*	*<u>noiose</u>*	*lezioni*	*di*	*Ferri*	*se*	*le*	*ricordano*	*tutti.*	*(unambiguous)*
	the	boring	classes	of	Ferri	remember			all	

b. *'Everybody remembers Ferri's classes, all of which were boring.' (nonrestrictive)*

c. *# 'Everybody remembers just those classes by Ferri that were boring.' (restrictive)*

(126) Italien (Cinque 2010 : 8), (Cinque 2014 : 4)

a.	*Le*	*lezioni*	*<u>noiose</u>*	*di*	*Ferri*	*se*	*le*	*ricordano*	*tutti.*	*(ambiguous)*
	the	classes	boring	of	Ferri	remember			all	

b. *'Everybody remembers Ferri's classes, all of which were boring.' (nonrestrictive)*
c. *'Everybody remembers just those classes by Ferri that were boring.' (restrictive)*

On peut résumer l'interprétation b de l'exemple (126) par la Figure 26, qui représente la sélection d'un ensemble qui est en réalité le même que l'autre ensemble – les deux cercles sont égaux car l'adjectif transmet une propriété inhérente au nom (interprétation non-restrictive, dite aussi « appositive »).

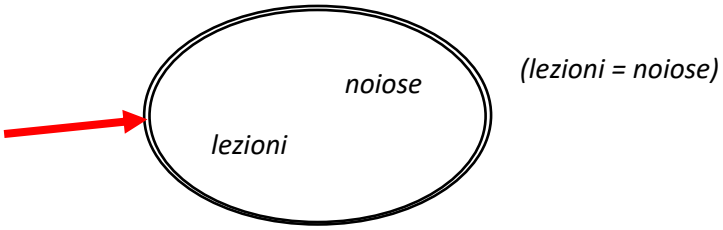


Figure 26 : Interprétation non-restrictive du syntagme (126)a,b

On peut résumer l'interprétation c par la Figure 27, qui représente la sélection d'un petit ensemble à l'intérieur d'un grand ensemble (interprétation restrictive, dite parfois déterminative).

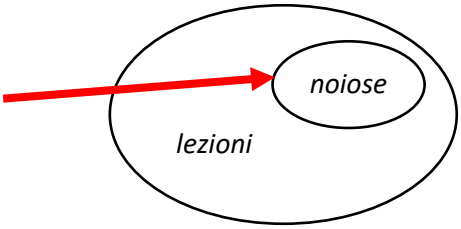


Figure 27 : Interprétation restrictive du syntagme (126)a,c

Lorsqu'un adjectif autorise deux interprétations, celle restrictive est la seule possible lorsque l'adjectif entre dans une structure polydéfinie (Lekakou et Szendrői 2012 : 131) :

(127) Grec (Lekakou et Szendrői 2012 : 131)

a.	*Gnorises*	*tin*	*orea*	*tin*	*tragudistria?*
	met.2SG	the	beautiful	the	singer
b.	'Did you met the beautiful singer?' (intersective only)				
	Gnorises	*tin*	*orea*	*tragudistria?*	
	met.2SG	the	beautiful	singer	
	'Did you met the beautiful singer?' (intersective and non-intersective)				
c.	*Gnorises*	*tin*	*orea?*		
	met.2SG	the	beautiful		
	'Did you met the beautiful one?' (intersective only)				

En (127)a, l'interprétation n'est qu'« intersective », selon le terme des autrices, et le référent est une personne qui danse de belle manière (et peut être par ailleurs physiquement laide). En (127)b, l'interprétation peut être intersective ou non (le référent étant alors une personne qui est physiquement belle, et qui peut par ailleurs danser de manière laide).

On sait que l'apposition étroite se caractérise par son interprétation restrictive : il existe une relation restrictive entre les deux sous-parties du syntagme. Elle est obligatoire, ce qui amène les chercheurs à parler de *contrainte de la polydéfinitude* : « *Greek polydefinites are unambiguously non-monotone anaphoric expressions: the discourse referent Y of a polydefinite is anaphoric to an* antecedent *discourse referent X, such that Y ⊂ X* » (Kolliakou 2004 : 273). Lekakou et Szendrői (2007 : 144) formulent la « contrainte du sous-ensemble propre » : « *In a close appositive, the denotation of one of the DPs must be a proper subset of the denotation of the other DP* ». Elle concerne les appositions étroites et donc, également les syntagmes polydéfinis. Un exemple comme (108)b, où Guillem est par définition un danseur (et un artiste), viole ce principe (de même que l'exemple (137) *infra*).

Il existe donc une asymétrie et une forme de hiérarchie entre les deux sous-parties du syntagme polydéfini, dans la mesure où l'un vient restreindre les référents possibles de l'autre. Celles-ci sont sémantiques et non syntaxiques. Elles sont à mettre sur le compte de l'ellipse du nom dans la partie adjectivale. La restrictivité implique également que l'adjectif soit « informatif », c'est-à-dire contienne une information différente de celle contenue dans la partie nominale, et différente des syntagmes précédents (de la même phrase ou non). Selon Lekakou et Szendrői (2007 : 149), c'est donc l'ellipse du nom qui rend l'interprétation de l'adjectif obligatoirement restrictive dans les appositions étroites en général et dans les syntagmes polydéfinis en particulier. Lekakou et Szendrői (2007 : 147) en déduisent qu'un syntagme polydéfini sera grammatical si « (*i) it introduces in the discourse a proper subset of a given set and (ii) that set has been explicitly mentioned (and is highly salient/accessible).* » Cependant la deuxième condition ne semble s'appliquer qu'aux syntagmes polydéfinis (Lekakou et Szendrői 2007 : 147). Le deuxième critère de grammaticalité s'explique par l'ellipse du nom, qui rend le syntagme informativement moins autonome. Il ne peut être énoncé que s'il a déjà mentionné l'ensemble d'entités auxquelles réfère le nom, l'ensemble de référents au sein duquel opérer l'opération de sélection.

Lorsque le syntagme présente une interprétation restrictive, c'est que le référent contraste avec un autre référent concurrent. Lorsque le syntagme présente une interprétation non-restrictive, c'est qu'une information supplémentaire, une propriété du référent est simplement délivrée, sans contraste sous-jacent. Du point de vue sémantique, l'une des deux sous-parties est nécessairement restrictive sur l'autre. C'est une sous-partie nominale, dans le cas des syntagmes DEF N DEF N, et la sous-partie adjectivale, dans le cas des syntagmes DEF A DEF N et DEF N DEF A, car elle modifie un nom nul (Lekakou et Karatsareas 2016 : 194). En effet, les adjectifs en contexte d'ellipse du nom sont obligatoirement interprétés restrictivement.

C'est à la restrictivité que l'on différencie, du point de vue sémantique, un syntagme monodéfini d'un syntagme polydéfini. C'est là que réside la différence de sens entre les deux types de syntagmes (Kolliakou 1998 : 15, Kolliakou 2004 : 270, Campos et Stavrou 2004 : 150, Lekakou et Szendrői 2009). Dans les syntagmes polydéfinis, l'adjectif est toujours un modifieur restrictif, tandis que dans les syntagmes monodéfinis, il peut être restrictif ou bien non-restrictif.

(128) Grec (Kolliakou 2004 : 270)

a. *O diefthindis dilose oti* ***i kali erevnites*** *tha eprepe na apolithun.*
'The director declared that the competent researchers should be fired.' (Two readings)

b. *O diefthindis dilose oti* ***i kali i erevnites*** *tha eprepe na apolithun.*
'The director declared that the competent researchers should be fired.' (Restrictive reading only)

Dans l'exemple (128)b, le syntagme nominal restreint les référents possibles aux *chercheurs compétents*, en sélectionnant seulement ceux-ci parmi l'ensemble des *chercheurs* (dont certains sont compétents, et d'autres non : les premiers seront licenciés, les seconds non). Dans l'exemple (128)a, le syntagme nominal peut recevoir cette interprétation, mais il peut également désigner un ensemble de *chercheurs* (qui se trouvent être tous compétents). Il n'est pas fait mention d'un autre ensemble de chercheurs, et ils ne sont pas opposés à des chercheurs non-compétents. Il n'est pas fait mention non plus de chercheurs qui ne se feraient pas licencier. L'adjectif renvoie à une propriété des chercheurs, mais celle-ci n'est pas décisive : il s'agit d'une information supplémentaire qui est donnée comme de façon annexe. On peut même aller plus loin : dans l'exemple (b), la compétence des chercheurs est interprétée comme la cause de leur renvoi – du moins c'est la seule cause qui peut être décelée dans la phrase ; dans l'exemple (a), aucune cause n'est donnée pour le renvoi et aucune n'est décelable dans la phrase. Dans la structure polydéfinie, l'adjectif se voit donc conférer une importance sémantique capitale au sein du syntagme et à l'échelle de l'énoncé. Il sélectionne un sous-ensemble parmi un ensemble de référents, et contraste donc ce sous-ensemble par rapport au reste de l'ensemble. Ce n'est pas le cas des structures monodéfinies, où l'adjectif revêt un sens simplement informatif et non sélectif.

Cependant, il convient de noter que tout le monde ne s'accorde pas sur l'interprétation à donner au syntagme polydéfini. Marinis (2003 : 190) partage l'interprétation de Kolliakou selon laquelle il doit toujours s'interpréter de façon restrictive. Kariaeva (2004) observait que l'ordre DEF N DEF A n'est pas correct dans les contextes où la partie nominale doit s'interpréter de façon restrictive par rapport à l'antécédent. Kyriakaki (2010 : 5-6) quant à elle, affirme qu'une interprétation non-restrictive est possible avec un syntagme polydéfini, en fonction de la syntaxe :

- si l'adjectif accompagné de son article défini est antéposé au nom, l'interprétation peut être seulement restrictive. Il apporte une information nouvelle.
- si l'adjectif accompagné de son article défini est postposé au nom, l'interprétation peut être restrictive ou non-restrictive.

En (129)a, une seule interprétation est possible : *préféré*, et non *cher*. L'interprétation ne peut être que restrictive, c'est-à-dire que la lecture est intersectionnelle et conjointe. La phrase sous-entend qu'il y a plusieurs *maris* et que l'un d'entre eux a la propriété d'être le *préféré*. Cependant si l'adjectif est postposé au nom, l'interprétation non-restrictive devient possible et l'interprétation est ambiguë.

(129) Grec (Kyriakaki 2010 : 5)

a.

O	*ayapimenos*	*mu,*	*o*	*andras*
The-MASC.NOM	beloved/favorite-MASC.NOM	my-CL.GEN	the-MASC.NOM	man-MASC.NOM

My favorite husband (# my beloved one, my husband)

b.	*O*	*andras*	*mu(,)*	*o*	*aɣapimenos*
	The-MASC.NOM	man-MASC.NOM	my-CL.GEN	the-MASC.NOM	beloved/favorite-MASC.NOM

'My favorite husband' ('My beloved husband')

En (129)b, le syntagme peut avoir deux sens : soit il n'y a qu'un *mari* et il a la propriété d'être *chéri*, soit il y a plusieurs *maris*, dont l'un a la propriété d'être *préféré*. Elle montre ainsi qu'en grec, l'interprétation restrictive ou non du syntagme est liée à l'ordre des constituants au sein de celui-ci. L'adjectif et son article fonctionnent de la même manière qu'une proposition relative, ou encore que les noms propres, les autres modifieurs introduits par un démonstratif ou un possessif. La place initiale et non marquée d'un tel modifieur à interprétation restrictive est postposée au nom.

Kyriakaki (2010 : 5-6) propose une analyse qui mette en lien la sémantique et l'ordre des constituants. Elle remet en question également le fait que l'ensemble des articles définis soit explétif, comme l'affirment Androutsopoulou (1995) ou Lekakou et Szendrői (2012). Selon elle, le fait qu'il accompagne obligatoirement les noms propres n'est pas un argument pour affirmer qu'il est explétif, et qu'il ne contribue pas à la définitude. Cela montre seulement qu'il est sous-spécifié en termes de définitude et n'encode pas toutes les valeurs possibles de celle-ci. Il est possible, par exemple, qu'il n'encode pas l'unicité, mais la familiarité (Kyriakaki 2010 : 4). C'est ici une différence notable entre le grec et l'anglais, qui n'autorise pas de structures polydéfinies.

L'autrice en déduit que les syntagmes à interprétation restrictive et ceux à interprétation non-restrictive sont structurellement différents. L'interprétation non-restrictive n'est possible qu'avec un adjectif postposé. C'est à mettre en lien avec l'information nouvelle et le focus que porte l'adjectif antéposé, qui force une interprétation restrictive. L'adjectif postposé porterait une information moins nouvelle, et moins de focus, d'où une possible interprétation non-restrictive. L'adjectif à interprétation restrictive est donc soit antéposé, soit postposé au nom. Un tel adjectif, antéposé au nom, porte un focus contrastif ou informatif. Il est dans ce cas porteur d'une information nouvelle. Postposé au nom, l'adjectif ne porte pas de focus et est le seul à permettre une interprétation non-restrictive. Il peut porter un focus s'il présente une pause et un changement intonationnel. Il est alors purement appositif, et non déterminatif – c'est-à-dire qu'il entre dans une apposition lâche et non étroite. Il apporte une information non-nécessaire à la reconnaissance du référent, un commentaire, une propriété dont la connaissance est facultative. Cela signifie donc qu'un adjectif postposé au nom peut présenter deux interprétations possibles : restrictive ou non-restrictive.

Selon Kyriakaki (2010 : 11), les noms propres et les noms dénombrables en apposition étroite partagent avec les adjectifs la caractéristique d'avoir une interprétation restrictive. Elle baptise « *polydefinite modifier* » un modifieur précédé d'un article défini qui sert à sélectionner un référent unique au sein d'un ensemble d'identités similaires (Kyriakaki 2010 : 18-19). Ce modifieur peut être adjectival (syntagme polydéfini « classique » avec un adjectif) ou nominal (apposition étroite « classique » avec nom propre et nom commun, ou nom propre à plusieurs éléments). Kyriakaki (2011 : 94) considère que le modifieur polydéfini a pour structure sous-jacente un DP contenant plusieurs DPs : une « matrice » et un (ou plusieurs) DP ayant la fonction de modifieur. Kyriakaki (2010 : 12) étend cette analyse aux démonstratifs, dans les syntagmes nominaux du type « *ce N* », qui ont en grec la structure DEF N DEM, à interprétation restrictive, ou DEM DEF N (avec accent sur le démonstratif), à interprétation restrictive et focalisée. La structure DEM DEF N sans accent sur le démonstratif est d'ailleurs agrammaticale. La différence entre tous ces types d'appositions étroites

est que le démonstratif est un pronom capable de sélectionner le référent à lui tout seul, et que le nom propre comme le nom dénombrable sont des têtes nominales et non adjectivales, structurellement autonomes. Cependant, tous occupent structurellement la même position, sont nécessaires au sens du syntagme donc à l'identification du référent, et peuvent porter un focus (Kyriakaki 2010 : 14).

Alexiadou (2014 : 40-41) contredit ces analyses, qualifiant de *restrictive* l'interprétation du syntagme avec adjectif postposé, et d'*émotive* l'interprétation du syntagme avec adjectif antéposé. Elle présente plusieurs cas ambigus, dont celui de l'adjectif *eksipnos*, intelligent. Dans le syntagme (130)a, il prend son sens littéral, de même que dans le syntagme (130)b, où il signifie « mon fils intelligent », par opposition à l'autre ou aux autres fils non intelligents. Une telle lecture contrastive est permise par la postposition de l'adjectif.

(130) Grec (Alexiadou 2014 : 41)

a.	*o*	*eksipnos*	*gios*	*mu*		D A N
	the	clever	son	my		
	my clever son					
b.	*o*	*gios*	*mu*	*o*	*eksipnos*	D N D A
	the	son	mine	the	clever	
c.	*o*	*eksipnos*	*o*	*gios*	*mu*	D A D N
	the	clever	the	son	mine	

L'antéposition, elle, ne permet pas de lecture restrictive ou contrastive, malgré la polydéfinitude. Le syntagme (130)c ne présuppose nullement l'existence d'autres fils que mon fils intelligent. Il n'est donc pas fait référence à celui de mes fils qui est intelligent. Le syntagme génère selon elle une interprétation « émotive », qui en vient à signifier le contraire, c'est-à-dire que mon fils n'est pas intelligent du tout. Lekakou et Szendrői (2012 : 109) confirment l'absence de contraste et de restrictivité, ainsi que l'interprétation « spéciale » du syntagme (130)c. On trouve le même phénomène avec l'adjectif *athoos*, innocent.

(131) Grec (Alexiadou 2014 : 40)

a.	*to*	*vrefos*	*to*	*athoo*
	the	baby	the	innocent
	'restrictive reading: the baby that is innocent'			
b.	*to*	*athoo*	*to*	*vrefos*
	the the	innocent	THE	baby
	'emotive reading: it actually has the opposite meaning, the not so innocent baby; in fact the person might not be a baby at all'			

2.4.3. … de l'adjectif : les exclus de la polydéfinitude

Certains adjectifs ne peuvent entrer dans une structure polydéfinie. C'est le cas des adjectifs **relationnels et ethniques** (Alexiadou et Wilder 1998 : 314, Alexiadou 2001 parlent d'adjectifs « **thématiques** »), mais aussi des adjectifs qui font partie d'un **nom propre** (Lekakou et Szendrői 2012 : 130) et des adjectifs sémantiquement **lexicalisés** et syntaxiquement **figés** avec le nom. Ainsi la polydéfinitude est-elle impossible dans le syntagme grec « l'énergie nucléaire », « la Maison-Blanche » ou « le Pôle Nord », littéralement « le pôle nordique » (Alexiadou et Wilder 1998 : 315),

ou encore « l'Océan Indien » (Campos et Stavrou 2004 : 163). En effet ces syntagmes forment une sorte de mot-composé où l'adjectif n'est pas tant un modifieur qu'une propriété intrinsèque de l'ensemble adjectif + nom.

- L'adjectif a sémantiquement fusionné avec le nom pour former un ensemble indécomposable et non modifiable, même en français : **l'énergie vraiment/très nucléaire.*
- Le mot composé ainsi formé semble intrinsèquement défini : **une énergie nucléaire, *quelques énergies nucléaires.*
- Ils ne peuvent pas non plus être en position d'attribut du sujet : **L'énergie est nucléaire*.
- Ils ne peuvent être substantivés ou supporter une élision du nom : **la nucléaire*.[1]
- La seule manière pour eux d'entrer dans une structure polydéfinie, c'est de rester uni avec le nom. Ils forment un bloc DEF A N syntaxiquement non modifiable, que peut éventuellement précéder un adjectif et un article, DEF A. L'ensemble DEF A DEF A N ainsi formé se structure DEF A [DEF A N], avec deux articles et deux adjectifs de nature très différente :

(132) Grec (Lekakou et Szendrői 2012 : 131)

a.	*o*	*diasimos*	*o*	*ekdotikos*	*ikos*
	the	famous	the	publishing	house
	'the famous publishing house'				
b.	**o*	*ekdotikos*	*o*	*diasimos*	*ikos*
	the	publishing	the	famous	house

On a vu dans la section 2.3.2, p. 164, que les adjectifs **non prédicatifs** sont exclus des syntagmes polydéfinis selon plusieurs auteurs (Alexiadou 2005 : 810). Mais tout dépend en fait du contexte. « *What is interesting to note is that manipulating the discourse context in the appropriate way can give us the desired effect even without an intersective adjective. The exceptional behaviour will arise whenever we can tamper with the potential of the noun denotation to be partitioned in disjoint subsets in a pragmatically plausible way.* » (Lekakou et Szendrői 2007 : 146). L'exemple (93) de l'adjectif *italiki*, italienne, répété ci-dessous, est considéré comme grammatical « *in appropriately manipulated discourse contexts* », par Lekakou et Szendrői (2007 : 138), par exemple si plusieurs invasions sont menées en même temps, dont une italienne que l'on veut distinguer des autres.

(93) Grec (Alexiadou et Wilder 1998 : 306)

i	*italiki*	*(*i)*	*isvoli*
the	Italian	the	invasion

Citons également les adjectifs *proin* (exemple (133), ancien, et *proighoumenos*, précédent. Ils sont jugés inacceptables en syntagme polydéfini par certains :

(133) Grec (Campos et Stavrou 2004 : 163), (Panagiotidis et Marinis 2011 : 42)

*	*o*	*proedros*	*o*	*proin*
	the	president	the	former

Mais ils sont jugés acceptables par Androutsopoulou (2001 : 191), Leu (2009 : 16), Lekakou et Szendrői (2007 : 146), Lekakou et Szendrői (2012 : 131-132),[2] dans le contexte où un locuteur

[1] On dit toutefois « le nucléaire » pour « l'énergie nucléaire », par exemple dans « prendre position contre le nucléaire ».
[2] C'est pourquoi Alexiadou (2014 : 33) reconnaît en note qu'il existe bel et bien une variation individuelle en ce qui concerne le type d'adjectif autorisé dans une structure polydéfinie.

corrigerait son interlocuteur ignorant la mort du premier ministre actuel et accentue prosodiquement l'adjectif (exemple (134). La dénotation du nom comprendrait alors deux référents possibles, le premier ministre actuel et le précédent.

(134) Grec (Lekakou et Szendrői 2012 : 132)

o	*PROIGUMENOS*	*o*	*prothipurgos*	*pethane.*
the	previous	the	prime.minister	died

'It is the previous prime minister that died.'

Selon Lekakou et Szendrői (2007), ce n'est pas la prédicativité le facteur pertinent, mais le potentiel de dénotation du nom, sa possibilité de rendre l'interprétation intersective et l'adjectif restrictif. Selon les autrices, la coïncidence entre les adjectifs qui peuvent apparaître dans les syntagmes polydéfinis et en position d'attribut du sujet n'est que partielle et accidentelle. Et même si l'adjectif n'est pas intersectif ni restrictif, le contexte peut pallier en fournissant l'ensemble des référents possibles où opérer une sélection.

On a vu dans la section précédente 2.4.2, p. 179, que les adjectifs **non-intersectifs** sont exclus des syntagmes polydéfinis selon beaucoup d'auteurs (Kolliakou 2004 : 264). Les adjectifs entraînant une interprétation non-intersective ne peuvent entrer que dans une construction monadique. En (135), l'adjectif *ipotithemeni*, prétendus, s'il entre dans une telle structure, doit sélectionner un sous-ensemble propre au groupe des terroristes, celui des prétendus. Mais *ipotithemeni* ne peut pas dénoter l'ensemble des terroristes, ni dénoter les entitées prétendues ; il peut seulement identifier des entités prétendues parmi les terroristes.

(135) Grec (Kolliakou 2004 : 264)

	i	*ipotithemeni*	*tromokrates*
	the	alleged	terrorists
*	*i*	*ipotithemeni*	*i*	*tromokrates*

(136)

	i	*apli*	*simptosi*
	the	mere	coincidence
*	*i*	*apli*	*i*	*simptosi*

On a vu également que les adjectifs **non-restrictifs** sont exclus des syntagmes polydéfinis selon beaucoup d'auteurs (Alexiadou 2005 : 810). L'absence de restriction pragmatique bloquerait la polydéfinitude :

(137) Grec (Kolliakou 2004 : 274)

a.

Idame	*tis*	*dilitiriodis*	*kobres.*
saw.1PL	the	poisonous	cobras

'We saw the poisonous cobras.'

b.

#Idame	*tis*	*dilitiriodis*	*tis*	*kobres.*
saw.1PL	the	poisonous	the	cobras

Le référent de (137) est à l'intersection des choses venimeuses et des cobras, mais l'adjectif venimeux n'exerce aucune restriction au sein des cobras, car tous les cobras sont venimeux et il n'existe pas de cobras non-venimeux : l'ensemble {cobras} et l'ensemble {cobras venimeux} se confondent. Kolliakou (2004 : 274) explique qu'un syntagme polydéfini est impossible avec la combinaison *cobras+venimeux*, car « *pragmatically non-restrictive terms cannot be polydefinites* ». En effet, la

réalité ne laisse pas d'autre interprétation possible que hiérarchique, où la totalité des cobras sont inclus dans les entités venimeuses : il n'y a pas d'interprétation conjointe possible, c'est-à-dire pas de restriction pragmatique possible.

2.4.4. … de l'adjectif : les exceptions

Chatzikyriakidis (2014, 2015) partage le point de vue général sur l'interprétation restrictive des syntagmes polydéfinis. Il présente toutefois une série d'autres exemples polydéfinis problématiques, puisque non-restrictifs.

Grec (Chatzikyriakidis 2015 : 13-14)

(138)	*O*	*eksipnos*	*o*	*aderfos*	*mu*	*pige*	*telika*			
	The	smart	the	brother	my	went	finally			
	'My wise-ass brother went in the end.'									
(139)	*Paok,*	*fere*	*mas*	*to*	*kipelo*	*ston*	*pirgo*	*ton*	*lefko*	
	Paok	bring	us	the	cup	to-the	tower	the	white	
	'Paok, bring the cup to the white tower'									
(140)	*Ta*	*megala*	*ta*	*ktiria*	*pu*	*se*	*elega*	*ina*	*aqfta*	*edo*
	The	big	the	buildings	that	you	told	are	these	here
	'The big buildings I was telling you about are these.'									
(141)	Context: *O Giorgos ixe tris mikres gates tis opies agapuse ke den apoxorizotan pote.*									
	Parola afta,		*mia*	*mera*	*i*	*mikres*	*i*	*gates*	*to*	*skasan*
	However		one	day	the	small	the	cats	it	ran
	'Georges had three small cats that he loved and never parted with them. However, one day the small cats ran away.'									
(142)	*Vgike*	*ekso*	*ston*	*krio*	*ton*	*kero.*				
	went	out	to-the	cold	the	weather				
	'S/He went out into the cold weather.'									

Notons que l'adjectif est antéposé dans quatre des cinq exemples (sauf (139). L'auteur établit un classement sémantique de ces exceptions :

- **Les syntagmes à connotation péjorative :** Dans l'exemple (138), il n'est pas dit que j'aie un autre frère. Chatzikyriakidis (2014 : 1) considère *eksipnos* comme ayant fonction d'*epithet*, « caractérisation », « étiquette », « formule » et non d'adjectif. Il entend par là une expression référentielle unique, où l'adjectif représente l'équivalent d'un titre apposé à un nom commun ou propre. Il est lié à lui de la même manière qu'un nom composé. Pour cette raison, il ne considère même pas un tel syntagme comme polydéfini. Non seulement l'interprétation du syntagme n'est pas restrictive, mais en outre sa sémantique n'est pas compositionnelle, elle n'est pas le résultat de l'association sémantique des termes *eksipnos* et *aderfos*. Le syntagme dans son ensemble présente une connotation péjorative : au lieu de « malin », il prend le sens de « qui prétend être

malin mais ne l'est pas ». Le syntagme fonctionnerait comme un mot composé exocentrique, impliquant un sens supplémentaire par rapport aux deux éléments qui le constituent. Cela rappelle les composés tels que *rouge-gorge*, qui n'est pas qu'une gorge rouge mais qui est aussi un oiseau (ici c'est le nom qui porte le sens supplémentaire), ou encore *grand-parent*, qui n'est pas qu'un parent grand à proprement parler, mais ancien (ici c'est l'adjectif qui porte le sens supplémentaire). Ici en grec, c'est l'adjectif qui porte le sens supplémentaire. L'auteur signale le même phénomène dans l'exemple (143), avec une apposition lâche où l'adjectif est postposé :

(143) Grec (Chatzikyriakidis 2014 : 1)

O	*aderfos*	*mu*	*pige*	*telika*	*o*	*eksipnos*
the	brother	my	went	finally	the	smart

'My wise-ass brother went in the end'

(Chatzikyriakidis 2014 : 3) parle de déviation par rapport au sens standard. Il mentionne d'autres adjectifs du même acabit : *anekdiigitos*, indescriptible, *terastios*, énorme, *kamenos*, brûlé (mais « fou » si antéposé en syntagme polydéfini étroit ou ou postposé en syntagme lâche). Notons que *eksipnos* a été analysé avec l'exemple (130) où, antéposé, il présentait une interprétation contrastive, mais, postposé, une interprétation « émotive » (Alexiadou 2014 : 40-41). Que la postposition de l'adjectif permette une lettre contrastive restrictive contredit l'exemple (143)…

- **Les noms propres composés :** Dans l'exemple (139), la tour, située à Thessalonique, n'est pas opposée à une autre tour qui serait d'une autre couleur – elle porte ce nom de manière historique (Figure 28). Ce monument est unique, l'interprétation n'est donc pas restrictive. Là aussi, le syntagme fonctionne comme un nom composé… à ceci près qu'il pourrait être un syntagme à ancienne interprétation restrictive (s'il existait jadis d'autres tours dans la ville, d'une autre couleur ou présentant une autre propriété), qui s'est figé comme nom propre et qui ne s'oppose plus, en synchronie, à un autre (ces autres tours n'existent plus). Notons qu'il ne s'agit pas là d'une propriété des noms propres. L'interprétation est restrictive lorsqu'une ambiguïté est possible, par exemple pour les anthroponymes – ce qui n'est pas le cas de la Tour Blanche, monument de Thessalonique, qui n'entre en opposition avec aucune autre tour. Il n'existe pas d'ensemble englobant toutes les tours au sein duquel on pourrait individuer une tour blanche. L'auteur constate que d'autres noms propres entrent moins bien dans une structure polydéfinie : le fleuve Jaune (Huang He ou, en grec, ο Κίτρινος Ποταμός, o Kítrinos Potamós), le Pont-Euxin

Figure 28 : La Tour Blanche (Thessalonique)

/ la Mer Noire (Ο Εύξεινος Πόντος, *o* Éfxeinos Póntos / ή Μαύρη Θάλασσα, *i* Mávri Thálassa) ou le Cap-Vert (Το Πράσινο Ακρωτήριο, *to P*rásino Akrotírio). Une explication serait le sens non compositionnel : le Cap-Vert est un État et un ensemble d'îles, mais n'est pas, au sens propre, un cap de couleur verte ; le fleuve Jaune est un fleuve et un concept, mais pas, littéralement, jaune). Au contraire, la Tour Blanche est bel et bien une tour de couleur blanche. La structure polydéfinie serait cependant grammaticalement possible dans un contexte opposant une mer *Euxin* (adjectif), c'est-à-dire hospitalière, à une mer *Axin* (adjectif), c'est-à-dire non-hospitalière. Il faudrait encore poser l'existence d'un Cap-*Rouge* et d'un Fleuve *Rose*. Dans ce cas, l'adjectif prendrait un sens littéral et le syntagme une sémantique compositionnelle. L'interprétation deviendrait restrictive.

- **Les syntagmes longs :** Dans l'exemple (140), les grands bâtiments ne sont pas opposés à d'autres qui auraient été précédemment mentionnés. L'auteur explique que plus le syntagme est long, plus il a de probabilité d'être interprété de manière non-restrictive. Ici, c'est la subordonnée relative qui vient « allonger » le syntagme – mais il donne d'autres exemples de modifieurs pouvant jouer le même rôle. Ils apportent des précisions sur le référent : son identification est assurée conjointement par l'ensemble des modifieurs qui portent sur le nom. L'équivalent 'court' de tels syntagmes, avec moins de modifieurs, est nécessairement interprété de manière restrictive. Dans le cadre de la théorie d'Ariel (1988, 1991), l'accessibilité du référent diffère d'un syntagme à l'autre et forme un continuum. Trois facteurs viennent l'influencer : l'*informativité* (la quantité d'informations donnée à propos du référent), la *rigidité* (dans quelle mesure le syntagme est discriminant et permet d'éliminer d'autres référents possibles) et la *taille phonologique* (la longueur du syntagme et la présence ou non d'un accent sur celui-ci). Il existe une corrélation entre ces facteurs : plus le syntagme est long, plus il a de chance d'être informatif, donc plus il est en mesure d'éliminer les référents similaires à celui visé. Ainsi les syntagmes polydéfinis peuvent-ils recevoir une interprétation non-restrictive si leurs référents sont très bas en termes d'accessibilité. Cette explication résoudrait le problème soulevé par Alexiadou (2014 : 41) :

  (144) Grec (Alexiadou 2014 : 41)

to	*prasino*	*to*	*kolovo*	*to*	*fidi*
the	green	the	tail-less	the	snake

  Toutefois ce syntagme ne présente que des adjectifs antéposés (structure DEF A DEF A DEF N). Elle propose de les analyser comme des cas de substantivisation.

- **Les syntagmes signalant un changement de topique :** Dans l'exemple (141), les petits chats peuvent parfaitement représenter l'ensemble des chats de Georges. Cela peut s'expliquer selon lui par le fait que la polydéfinitude constitue le marqueur d'un changement de topique (ou de thème) dans la structure de l'information de la phrase. La locution adverbiale *un jour* signale elle aussi un tel changement. Les *petits chats*, qui auparavant étaient le rhème, constituent le thème de la nouvelle phrase. Le fait de participer d'un changement de topique rend possible l'interprétation non-restrictive. Cela peut sembler étonnant, dans la mesure où l'on a vu que la structure polydéfinie participe à la focalisation.

- **Les cas non résolus :** Dans l'exemple (142), le froid n'est pas opposé à un temps chaud. Chatzikyriakidis (2016) cite également l'expression *ta kala ta pedia, good lads*, les gars / les amis, il ne parvient pas à expliquer l'existence d'une interprétation non-restrictive. Il ne parvient pas non plus à déterminer la différence pragmatique et sémantique avec sa contrepartie monodéfinie. Panagiotidis et Marinis (2011 : 10) mentionnent l'exemple (142) et comme l'expression *ta pedia ta kala* (même sens, mais traduction par « les enfants sages », avec adjectif postposé). Ils précisent seulement que cette dernière s'emploie dans la langue quotidienne spontanée (*run-of-the-mill usage*).

Signalons enfin deux cas particuliers, celui des adjectifs d'identités : l'adjectif « même », ίδιος, *iðios,* peut entrer dans une structure polydéfinie. Il a alors le sens d'un renforcement, similaire au français « elle-même », « nous-mêmes ».

- il peut être utilisé comme un nom (ο ίδιος, *o iðios*, le même, de structure DEF N),
- s'il est utilisé comme un adjectif, il ne peut entrer dans un syntagme monodéfini (DEF A N),
- il peut seulement être utilisé comme un adjectif articulé dans un syntagme polydéfini (DEF A DEF N, beaucoup plus probablement que DEF N DEF A).

Selon Kyriakaki (2011 : 24), cet usage étonnant de ίδιος constitue une preuve qu'il fonctionne comme un élément nominalisé.

L'adjectif « autre », άλλος, *allos,* peut lui aussi entrer dans une structure polydéfinie, alors même qu'il ne peut exercer la fonction d'attribut du sujet – ce qui met là aussi à mal la comparaison entre adjectif polydéfini et prédication (Lekakou et Szendrői 2012 : 133) :

(145) Grec (Lekakou et Szendrői 2012 : 133)

a.	*i*	*ali*	*i*	*singrafeas*
	the	other	the	author
	'the other author'			
b.	**I*	*singrafeas*	*ine*	*ali*
	the	author	is	other

En effet, sa distribution est différente de celle d'un adjectif car il peut précéder le déterminant lorsque celui-ci est un article indéfini (*un autre livre de X* prononcé « autre un livre de X »). On le trouve donc souvent décrit comme un pronom (Marinis 2003 : 185). Comment catégoriser cet adjectif ? Le sens d'un tel adjectif dépend non seulement de celui du nom, mais aussi du co-texte et du contexte d'énonciation. Il a le même sens que l'adjectif français *différent* interprété de manière externe[1], c'est-à-dire dénotant une propriété non intrinsèque du référent, contrastive de manière binaire (*x ~ autre*). Cette propriété consiste, pour le référent, à être différente d'un référent déjà mentionné, saillant dans le contexte. *Autre* peut également avoir le sens de *restant* ou de *second*. On peut alors considérer l'interprétation du syntagme nominal « l'autre auteur » comme intersective en apparence subsective.[2]

[1] *Différent* a une deuxième interprétation, qualifiée d'interne, dans *Il y a différentes pommes de terre sur la table*. L'adjectif est ici plus difficile à interpréter car il ne dénote pas une propriété du référent : il ne serait pas, alors, intersectif. *Vaver* romani (et *autre* français) ne présente pas, dans notre corpus, cette interprétation interne de *différent*.

[2] Je tiens à remercier Marcin Morzycki pour avoir discuté cette question avec moi.

2.5. Analyse informationnelle

Plusieurs auteurs ont senti l'existence d'un focus en rapport avec le syntagme polydéfini. L'interprétation restrictive de l'adjectif serait par exemple due au focus qu'il reçoit. Le syntagme polydéfini existerait donc pour des raisons de structure de l'information.

- Le focus est-il situé sur la totalité du syntagme polydéfini ou sur sa sous-partie adjectivale ?
- Est-il un focus syntaxique et/ou un focus prosodique (via l'accent mis sur l'adjectif à l'oral) ?
- A quel niveau de la structure syntaxique FocP se situe-t-il et y a-t-il un mouvement ?
- Du point de vue informationnel, de quel type de focus s'agit-il ?

1) Kolliakou (2004 : 275) nie que la sous-partie adjectivale puisse comporter un focus subordonné, intégré dans le syntagme nominal constituant par exemple le topique de la phrase. L'inexistence d'un focus pragmatique « subordonné » à l'intérieur d'un groupe de statut topique n'est pas avérée dans les langues du monde, ce qui infirme selon elle une telle analyse pour le grec. L'autrice montre que le syntagme polydéfini peut faire partie soit du focus (dans des phrases comme (70)d, répétée ci-dessous), soit du topique de la phrase (dans des phrases comme (146). L'accent (*pitch*) que comporte l'adjectif *mikres*, petits, dans l'exemple (146), est celui d'un topique, selon elle incompatible avec l'analyse d'un tel élément en tant que focus.

(70) Grec (Lekakou et Szendrői 2014 : 219, inspiré de Kolliakou 2004 : 269)

a. Speaker A:	*Ti*	*pires*	*tu*	*Janni*	*ja*	*ta*	*christujena ?*
	what	took.2SG	the.GEN	John	for	the	Christmas

'What did you get John for Christmas?'

b. Speaker B:	*(Tu*	*pira)*	*tin*	*asimenia*	*pena.*
	him.GEN	took.1SG	the	Silver	Pen

'(I got him) the silver.'

c. Speaker A:	*Ti*	*pires*	*tis*	*Marias ?*
	what	took.2SG	the.GEN	Maria

'What did you get for Maria?'

d. Speaker B:	*(Tis*	*pira)*	*tin*	*pena*	*ti*	*chrisi.*
	her.GEN	took.1SG	the	pen	the	golden

'(I got her) the golden pen.'

(146) Grec (Kolliakou 2004 : 275)

[[I mikres] $_{FOC-sub}$ [i gates] $_{TOP-sub}$] $_{TOP}$ [ine eksipnes] $_{FOC}$

the young the cats are intelligent; 'the YOUNG cats are intelligent'

Selon Szendrői (2010 : 873), Kolliakou (2004) montre en fait que la différence pragmatique entre syntagme monodéfini et polydéfini ne consiste pas à focaliser l'adjectif, mais à désaccentuer le nom.

2) Selon Campos et Stavrou (2004 : 141-142) et Alexiadou (2005 : 809), les syntagmes polydéfinis impliquent un contraste et sont organisés autour de la distinction présupposition-focus. La présence

du déterminant adjectival réduit l'ensemble qui constitue la dénotation du nom (Alexiadou 2005 : 809). Campos et Stavrou (2004 : 147) expliquent le rôle important de l'accent dans l'interprétation[1] :

> « *If the adjective precedes the noun, it is contrastively stressed and the adjective is interpreted under narrow focus, with emphatic stress being assigned to it by the Emphatic/Contrastive Stress Rule.* […] *Thus, we see that in a monadic construction, depending on context, either the adjective or the noun may carry contrastive stress, whereas in a polydefinite the adjective is always the one to be stressed, the noun providing only contextualized/old information*. » (Campos et Stavrou 2004 : 147)

Selon Kolliakou (2004), Campos et Stavrou (2004) et Marchis et Alexiadou (2009), les deux sous-parties de la structure ne s'interprètent pas de la même manière et impliquent un contraste – comme en roumain. Elles s'organisent autour de l'opposition présupposition ~ focus : la sous-partie nominale constituant la présupposition (ou contexte), la sous-partie adjectivale constituant le focus.

Selon Cinque (2010 : 107) également, les syntagmes polydéfinis impliquent un contraste, ce qui est toujours le cas des DP contenant une ellipse du nom. Comme Cornilescu et Nicolae (2012 : 1071), il considère qu'un syntagme avec ellipse prend par la même occasion un focus contrastif (« *for a successful subdeletion there must be a contrast* » (Giannakidou et Stavrou 1999 : 304).

Alexiadou (2014) parle de *premier* élément et de *second* élément, le premier constituant le focus, le second la présupposition. La partie nominale du syntagme constitue le contexte, ou la présupposition ; la partie adjectivale constitue le focus. L'article défini « supplémentaire » d'une telle structure assume la même fonction qu'un focus contrastif placé sur l'adjectif d'un syntagme monodéfini.[2] Le tout a pour effet de contraster et de sélectionner un référent particulier parmi ceux dénotés par le nom. Le nom n'étant pas accentué, il constitue le contexte, l'ensemble de référents parmi lesquels effectuer la sélection.

Szendrői (2010 : section 1) s'inscrit en faux contre cette théorie : selon elle, les notions de topique et de focus, de *comment*, commentaire et de *background*, arrière-plan, sont des notions intrinsèquement propositionnelles. Elles ne peuvent donc pas s'appliquer au DP, syntagme déterminatif. La seule notion qui n'est pas intrinsèquement propositionnelle et qui est liée au discours, est la notion de *givenness*, donné, plus exactement l'anaphoricité en discours (Szendrői 2010 : 869).

3) Selon Lekakou et Szendrői (2007 : 151), la partie adjectivale d'un syntagme polydéfini n'est pas nécessairement sous focus. Cette sous-partie ne peut pas être anaphorique. Mais il ne s'ensuit pas nécessairement qu'elle porte un focus interne, du point de vue de la structure de l'information. Les autrices suivent Reinhart (2006) et séparent le statut de la structure de l'information, considérées comme des notions « orthogonales ». La partie adjectivale de la construction ne peut être ni donnée ni anaphorique, mais ne doit pas pour autant être focalisée. Elles s'appuient pour cela sur l'exemple (147)a, où la polydéfinitude constitue le topique (contrastif) de la phrase, en l'absence de tout focus contrastif, puisque la phrase peut se poursuivre en « tels que l'étaient les anciens ». Elles font en effet

[1] Cette affirmation sera contredite par Lekakou et Szendrői (2012).

[2] Cette affirmation entre en contradiction partielle avec Androutsopoulou (1995 : 182), selon laquelle les adjectifs antéposés au nom ne reçoivent pas de focus.

une distinction entre focus et topique contrastif. Elles comparent d'ailleurs cet exemple avec (147)b, où l'adjectif du syntagme polydéfini reçoit un accent prosodique. Il y a bel et bien un focus contrastif, et la poursuite de la phrase avec ces mots n'est pas possible. C'est l'accent qui provoque les effets du focus contrastif, dont l'interprétation exhaustive, et non la structure polydéfinie.

(147) Grec (Lekakou et Szendrői 2007 : 151)

a.	*O*	*Yannis*	*taise*	*ta*	*zoa.*	*I*	*mikres*	*i*	*gates*	*itan*
	The	Yannis	fed	the	animals.	The	young	the	cats	were
	pinasmenes	*opos*	*episis*	*ke*	*i*	*megales*	*(i*	*gates).*		
	hungry	as	also	and	the	old	(the	cats).		
	'Yiannis fed the animals. The young cats were hungry, as were the old ones.'									
b.	*O*	*Yannis*	*taise*	*ta*	*zoa.*	*I*	*MIKRES*	*i*	*gates*	*itan*
	The	Yannis	fed	the	animals.	The	YOUNG	the	cats	were
	pinasmenes,	*#*	*opos*	*episis*	*ke*	*i*	*megales*	*(i*	*gates).*	
	hungry		as	also	and	the	old	(the	cats).	

Lekakou et Szendrői (2012 : 134-135) approfondissent la question et montrent que c'est la proéminence prosodique, plus précisément l'accent mis sur l'adjectif, qui provoque l'effet de contraste associé habituellement avec le focus, et non la syntaxe polydéfinie. Du point de vue structurel, il serait impossible d'introduire un FocP à l'intérieur d'un DP.

4) Selon Szendrői (2010 : 870), le caractère donné est intimement lié à la désaccentuation prosodique, sur les constituants qui devraient normalement être accentués. Dans le cas des DP portant l'accent du côté droit et présentant un ordre ADJ N, l'accent porte donc sur le nom (ordre A N). Si le syntagme est donné, on peut donc désaccentuer le nom et accentuer l'adjectif (ordre A N), ou bien inverser l'ordre des mots et conserver l'accent du côté droit (ordre N A). Elle explique que le phénomène de polydéfinitude en grec repose non sur un phénomène de focus, mais sur un phénomène de donné, d'anaphoricité. On ne peut construire de syntagme polydéfini en mettant sous contraste la partie adjectivale, à moins que la partie nominale ne soit donnée. Il est plus important que la partie nominale soit donnée que la partie adjectivale ne soit accentuée ou focalisée (Szendrői 2010 : 873-874). Il est en effet impossible de trouver un syntagme polydéfini si la partie nominale n'est pas donnée – et ce, même si la partie adjectivale est accentuée et présente un focus contrastif. Seule une construction monodéfinie est possible. Elle l'illustre par l'exemple (148) :

(148) Grec (Szendrői 2010 : 874)

O	*Janis*	*ithele*	*ena*	*metaforiko meso*		*ja*	*tis*	*diakopes*	*tu.*
The	Janis	wanted	a	means-of-transport		for	the	holidays	his.
Pije	*sto*	*garage tu patera tu.*		*Epidi*	*to*	*KOKINO*	*(*to)*	*aftokinito*	
Went	to.the	father's garage.		As	the	red	the	car	
ihe	*idi*	*pulithi*	*pire*	*to*	*ble*	*((to)*	*aftokinito).*		
was	already	taken	he took	the	blue	the	car.		

'Janis wanted a vehicle for his holidays. He went to his father's garage. As the red car was already taken, he took the BLUE car.'

Il est également impossible de trouver un syntagme polydéfini s'il n'y a qu'un contraste d'adjectifs. Seule une construction monodéfinie est possible. Elle l'illustre par l'exemple (149) :

(149) Grec (Szendrői 2010 : 874)

Anigo	*tin*	*tileorasi*	*ke*	*ti*		*vlepo?*	*Ton*	AMERIKANO
switch.on.1SG	the	television	and	what		see.1SG?	The	American
*(*ton)*	*proedro*	*na*	*sinomili*	*me*	*ton*	*IRANO*	*(*/??ton)*	*proedro.*
the	president	subj	talk.with	with	the	Iranian	the	president.

"I switch on the telly and what do I see? The American president is talking to the Iranian president.'

Il y a un effet de focus sur la partie adjectivale, certes, mais celui-ci est dû à la présence d'une élision du nom, plutôt que *DP-internal focus fronting*, une mise en avant du focus interne au DP. Même dans les exemples construits, où la partie nominale n'est pas donnée mais où la partie adjectivale est focalisée et mise sous contraste, on ne peut trouver de syntagme polydéfini. Cela confirme la théorie de Kolliakou (2004), selon laquelle l'apport pragmatique du syntagme polydéfini est la désaccentuation de la partie nominale et son statut donné, plutôt que la focalisation de la partie adjectivale.

5) Une autre analyse est proposée par Kyriakaki (2010 : 5-6), selon laquelle les interprétations restrictive et non-restrictive ont une position structurale différente (cf. section 2.4.2, p. 179). La première est possible en position anté- ou postnominale. La seconde n'est possible qu'en position postnominale – qu'elle reçoive ou non un focus. Ceci est dû, selon elle, aux propriétés appositives de l'emplacement postnominal, à son caractère de « supplément » ou de « commentaire », qui ne fait pas partie de l'information principale de la partie nominale de la structure. En d'autres termes, l'adjectif n'a pas la même interprétation selon sa position. Lorsqu'il est postposé au sein du syntagme polydéfini, il peut déclencher une interprétation restrictive ou non-restrictive. Lorsqu'il est antéposé, l'interprétation ne peut être que restrictive et l'adjectif est dans une position de focus, où l'information est nouvelle, voire nouvelle et contrastive. Ainsi, c'est la place de l'adjectif et la structure de l'information qui détermineraient le type de lecture, restrictive ou non, du syntagme.

C'est le focus qui force l'interprétation restrictive. Or il existe deux types de focus : informatif et contrastif Kyriakaki (2010 : 5-6). L'exemple (150) peut se lire de deux façons, en raison de l'antéposition de l'adjectif : de façon informative (sans contraste, l'adjectif donne simplement une propriété du référent), ou de façon contrastive (l'adjectif oppose une propriété à une autre, avec la signification « c'est de la robe rouge dont j'ai besoin »). Au contraire, lorsqu'il est postposé, l'adjectif ne reçoit pas de focus (exemple (151)a), à moins d'introduire une pause, ce qui le rend non-restrictif (exemple (151)b).

(150) Grec (Kyriakaki 2010 : 6)

Xriazome	*[TO*	*KOKINO*	*to*	*forema].*
need-1SG	the-NEU	red-NEU	the-NEU	dress-NEU

'I need the RED dress.'

(151) a. *Xriazome* *to* *forema* *#(TO* *KOKINO).*
need-1SG the-NEU dress-NEU the-NEU red-NEU
'I need the red dress.'

b. *Xriazome* *to* *forema* – *TO* *KOKINO.*
need-1SG the-NEU dress-NEU the-NEU red-NEU
'I need the dress – the red one.'

Mais l'ordre des constituants détermine aussi l'interprétation à l'échelle de l'énoncé. Lorsque l'objet est antéposé au verbe (exemple (152), en position focalisée, le focus contrastif sur l'adjectif est obligatoire (exemple (152)a, excluant (152)b sans accent sur l'adjectif). L'antéposition de l'adjectif est elle aussi obligatoire, d'où l'agrammaticalité de (153) présentant un adjectif postposé – même avec un focus contrastif portant sur le nom.

(152) Grec (Kyriakaki 2010 : 6)

a. *[TO* *KOKINO* *to* *forema]* *xriazome.*
the-NEU red-NEU the-NEU dress-NEU need-1SG
'I need the RED dress.' (or '*it is the red dress I need*')

b. *#To* *kokino* *to* *forema* *xriazome.*
the-NEU red-NEU the-NEU dress-NEU need-1SG
'I need the red dress.'

(153) # *TO* *FOREMA/to forema* *to* *kokkino* *xriazome.*
The-NEU dress-NEU the-NEU red-NEU need-1SG
'I need the RED dress.' (or '*it is the red dress I need*')

On remarque que seul l'adjectif antéposé peut recevoir un accent prosodique. Contrairement à la position anténominale de l'adjectif, qui permet le focus et l'interprétation restrictive, celle postnominale ne permet pas le focus de l'adjectif, à moins d'une interprétation non-restrictive. Contrairement à la position antéverbale du syntagme polydéfini, celle postverbale permet le focus de l'adjectif, à condition qu'il soit antéposé. Le focus est alors soit informatif, soit contrastif.

2.6. Bilan

2.6.1. La polydéfinitude en grec

La structure polydéfinie est constituée d'un nom précédé d'un article défini et d'au moins un adjectif précédé d'un article défini.

- Il ne semble exister aucune contrainte quant au **type de nom**. Une fois en syntagme polydéfini, le nom ne peut porter l'accent.
- Il ne semble exister aucune contrainte d'ordre morphologique quant au **type d'adjectif**. Mais il existe une contrainte syntaxique : l'adjectif doit être prédicatif, c'est-à-dire pouvoir jouer la fonction d'attribut du sujet. Nous avons vu que ce point est toutefois débattu et qu'il y aurait des exceptions.

Il existe également une contrainte sémantique : l'adjectif doit être restrictif et impliquer un contraste sémantique. La partie adjectivale du syntagme doit restreindre la dénotation du nom, c'est-à-dire dénoter un sous-ensemble au sein de l'ensemble de référents dénotés par le nom. Les adjectifs relationnels sont exclus, de même que ceux qui font partie d'un nom propre, d'un mot composé ou qui sont lexicalisés et figés avec le nom – mais il existe là aussi des exceptions. L'adjectif doit déclencher une lecture intersective : l'interprétation du syntagme nominal doit être conjointe. Si plusieurs interprétations sont possibles dans un syntagme monodéfini, seule l'interprétation intersective subsiste dans le syntagme polydéfini correspondant. Nous avons vu là aussi qu'il existait des exceptions, avec des syntagmes polydéfinis non-restrictifs, non contrastifs et non-intersectifs.

- Il existe une contrainte quant au **type de déterminant** : il doit nécessairement être un article défini. L'article défini, ou numéral « un » grammaticalisé, est exclu. Le syntagme peut éventuellement être déterminé par un (clitique) possessif ou démonstratif supplémentaire. Le statut de l'article défini a été discuté : marque d'accord en définitude de l'adjectif, outil de restriction de la référence, cas de surdétermination, c'est-à-dire de détermination multiple avec une notion d'excès, porteur du trait de définitude ou article explétif (théorie du *Def-D-split*). Par exemple Lekakou et Szendrői (2014 : 213) ou Lekakou et Karatsareas (2016 : 199) considèrent l'article défini comme explétif. Les syntagmes polydéfinis ne le sont donc que du point de vue *morpho-syntaxique*, en raison des multiples instances de l'article défini. Mais ils ne le sont pas du point de vue *sémantique*, car le second article défini est explétif. Au contraire, selon Kyriakaki (2011 : 134), la fonction du syntagme polydéfini est de signifier une restriction du nom-tête grâce à l'article défini supplémentaire. On peut en déduire que ce dernier a un impact sémantique ; il n'est pas explétif mais est porteur du trait de définitude.

Sur le plan **morphologique**, si le syntagme polydéfini est fléchi, on observe une réplication de la désinence casuelle sur l'adjectif et son article défini. Si une préposition introduit le syntagme, elle est également répliquée en tête de la partie adjectivale. Sur le plan **syntaxique**, l'ordre des mots est plus souple dans un syntagme polydéfini que monodéfini. Ce dernier n'accepte que des adjectifs antéposés (non articulés). Le premier accepte tant les adjectifs antéposés (articulés ou non) que les adjectifs postposés (articulés). La position de l'adjectif épithète par rapport au nom dépend d'un certain nombre de facteurs sémantiques : la définitude, le contexte, le sens et le type d'adjectif. La postposition de l'adjectif correspondrait à une propriété temporaire du référent, au sens propre de l'adjectif, à une emphase sur l'adjectif. L'antéposition de l'adjectif permettrait un focus informatif ou contrastif, tandis que sa postposition ne permettrait pas de focus sur l'adjectif – à moins de transformer le syntagme en polydéfini lâche à interprétation non-restrictive.

Sur le plan **structurel**, l'adjectif serait le prédicat d'une proposition relative sous-jacente, réduite, avec ou sans ellipse du nom, ou bien le modifieur d'un nom élidé. Les premières théories postulent l'existence d'une hiérarchie entre les deux sous-parties du syntagme. La dernière postule leur égalité et leur symétrie – la partie adjectivale n'étant pas sous focus puisque seul l'accent prosodique provoquerait le focus contrastif. La partie adjectivale serait un cas de modification indirecte ou bien d'apposition étroite à interprétation restrictive. Il pourrait exister des syntagmes polydéfinis lâches où l'adjectif apporterait une information non essentielle et non-restrictive. Sur le plan **informationnel**, le syntagme polydéfini peut faire partie du topique ou du focus de la phrase. En interne, il s'organiserait autour de la distinction présupposition-focus, la partie nominale constituant

la présupposition et la partie adjectivale le focus. L'interprétation restrictive de cette dernière est soit due au seul focus (syntaxique) qu'il reçoit par la structure polydéfinie, soit à un accent prosodique s'ajoutant à la structure polydéfinie (focus prosodique + syntaxique). L'emplacement postnominal aurait des propriétés appositives, un caractère de « commentaire » de l'information principale contenue dans la partie nominale.

2.6.2. Problématique

La comparaison entre le romani et le grec devrait être très fructueuse car le phénomène est très semblable dans les deux langues. Des chercheurs comme Lekakou et Szendrői (2007 : 152) se sont étonnés que l'on ne trouve de syntagmes polydéfinis qu'en grec : ce n'est donc peut-être pas le cas. Dans ce travail, nous nous efforcerons de répondre aux questions suivantes (Figure 87) :

Questions structurelles

1) Quels éléments peuvent entrer dans un syntagme polydéfini romani ?
2) Dans quel ordre les trouve-t-on ? Existe-t-il des structures DEF A DEF N ?
3) L'ordre des mots est-il plus souple en syntagme poly- que monodéfini ?
4) Le romani présente-t-il la même souplesse que le grec ? Lorsqu'une expression polydéfinie contient plusieurs adjectifs, trouve-t-on les mêmes contraintes d'ordre qu'en grec ?
5) L'adjectif d'un syntagme polydéfini peut-il avoir son propre complément ?
6) L'emplacement postnominal est-il la zone des « options » ? Quel est le degré de nominalité des éléments qui s'y trouvent ?
7) Existe-t-il une restriction quant au type de nom pouvant entrer dans un syntagme polydéfini ?
8) Existe-t-il des contraintes morphologiques ou étymologiques quant au type de l'adjectif ?
9) Existe-t-il des contraintes syntaxiques ? L'adjectif doit-il être prédicatif ?
10) Existe-t-il des contraintes sémantiques ? Doit-il être restrictif ? Certains adjectifs sont-ils exclus, comme les adjectifs relationnels ou ceux lexicalisés et figés avec le nom ?
11) Existe-t-il des contraintes quant au type de déterminant ? Peut-il y avoir répétition d'autre chose que de l'article défini, notamment l'article indéfini ?
12) L'article défini du romani est-il explétif ?
13) Peut-il y avoir répétition du marquage de cas et de la préposition ?
14) La partie adjectivale est-elle un cas de modification directe ou indirecte ?
15) Elle-t-elle un cas d'apposition étroite ?

16) Existe-t-il des syntagmes polydéfinis lâches ? Leur interprétation est-elle non-restrictive ?
17) Y a-t-il un lien entre polydéfinitude et ordre des constituants nominaux ? Lequel ?

Questions sémantiques

1) Existe-t-il une nuance de sens entre syntagme monodéfini et syntagme polydéfini ?
2) *Poly*définitude implique-t-elle *sur*définitude ?
3) Les syntagmes monodéfinis sont-ils restrictifs, tandis que les syntagmes polydéfinis peuvent être restrictifs ou non-restrictifs ?
4) Si l'adjectif est antéposé au nom, l'interprétation est-elle seulement restrictive ? Si l'adjectif est postposé au nom, l'interprétation peut-elle être restrictive ou non-restrictive ?

Questions fonctionnelles

1) La polydéfinitude est-elle optionnelle en romani comme en grec ? Si oui, qu'est-ce qui provoque le choix du locuteur ?
2) S'agit-il d'une variation optionnelle de type *syntactic doubling* (Barbiers 2013 : 7) ?
3) Quelle est la fonction des éléments situés dans l'emplacement postnominal (épithète, apposition, *afterthought*…) ?
4) Le syntagme polydéfini implique-t-il un contraste ?
5) Quel est le statut de l'information qu'il porte ?
6) Fait-il plutôt partie du topique ou du focus de la phrase ?
7) Implique-t-il un focus sur l'adjectif ? Si oui, de quel type ?
8) Le focus est-il situé sur le la totalité du syntagme polydéfini ou sur sa partie adjectivale ? Dans le deuxième cas, la partie nominale constitue-t-elle la présupposition ?

Questions socio-linguistiques

1) Tous les locuteurs albanais du romani emploient-ils la structure polydéfinie ?
2) Observe-t-on une différence de profil socio-linguistique parmi les locuteurs employant cette structure : différence de genre, d'âge ou de niveau d'éducation ?
3) Quelles variétés présentent le plus cette structure ?
4) Les locuteurs qui emploient cette structure peuvent-ils être influencés par la connaissance d'une ou plusieurs langues étrangères, et si oui, lesquelles ?
5) Quel est le profil « type » de la personne employant le plus volontiers la structure polydéfinie ?
6) Quel est le profil « type » de la personne postposant le plus volontiers l'adjectif ?

7) Quel est le profil « type » de la personne antéposant le plus volontiers l'adjectif ?

Questions diachroniques

1) À quand remonte cette structure ? Est-il possible de la dater ?
2) D'où provient-elle ? S'agit-il d'une innovation interne ou d'un « emprunt » ?
3) Quel a été le processus d'acquisition de la structure ?
4) Existe-t-elle dans d'autres variétés du romani ? Constate-t-on des différences ?
5) La syntaxe du syntagme nominal est-elle stable en romani ?
6) La syntaxe du syntagme nominal est-elle stable en romani d'Albanie ?

Questions typologiques

1) La polydéfinitude existe-t-elle dans d'autres langues que le romani et le grec ?
2) Quels sont les différents « types » de polydéfinitude dans les langues d'Europe ?
3) Où se situe le romani dans cette typologie ?
4) Comment nommer au mieux la structure qui fait l'objet de la présente étude ?

Figure 29 : Problématique

Chapitre IV : Analyse du corpus spontané

Cette partie concerne les données qui forment le corpus « spontané ». Après avoir présenté le corpus et l'usage de la langue dans la famille qui m'a accueillie, on analysera qualitativement une cinquantaine d'exemples. Ceux-ci sont classés en fonction du type de structure qu'ils comportent : syntagme polydéfini (adjectif postposé, attendu comme le plus courant, et adjectif antéposé) puis syntagme monodéfini (adjectif antéposé, attendu comme le plus courant, et adjectif postposé). Un certain nombre de syntagmes polydéfini lâches ayant été relevé, on a également analysé ces structures présentant une rupture au sein du syntagme nominal : pause ou insertion lexicale. Elles sont présentées en suivant le même ordre : syntagme polydéfini avec rupture interne (à adjectif postposé) puis syntagme monodéfini avec rupture interne (adjectif antéposé et adjectif postposé).

1. Méthodologie

1.1. Élaboration du corpus spontané

Nous entendons par corpus spontané les enregistrements réalisés lorsque les locuteurs s'exprimaient de façon « libre » – en dehors d'un contexte de tâche à réaliser (cf. section 3.4, p. 53). Nous choisissons d'y inclure les conversations libres, les entretiens dirigés et semi-dirigés (visant à récolter des métadonnées, un récit de vie, un conte, une recette de cuisine…). Certains extraits sont issus de l'observation participante. Nous pensons en effet que l'étude des phénomènes liés à la structure de l'information est plus efficace si elle se fait sur un corpus d'énoncés tirés de la réalité plutôt que sur de l'élicitation ou des énoncés créés *ad hoc* – même si ces derniers peuvent parfois mieux satisfaire les conditions théoriques qui sont développées.

C'est pourquoi nous travaillerons sur des énoncés prononcés spontanément et dans divers contextes. Le contenu informatif et pragmatique des données est donc hautement variable et imprévisible, il est donc ni comparable ni systématique : il se prête à une analyse qualitative mais non quantitative. La question se pose alors de savoir combien de temps et de types d'enregistrements sont requis pour infirmer ou confirmer une hypothèse, donc pour établir une description fiable et complète des structures recherchées. En outre, le paradoxe de l'observateur induit entre autres une variation stylistique : certains locuteurs ont modifié ponctuellement leur lexique, leur syntaxe, voire adopté une

attitude professorale envers moi ou les autres participants de la conversation[1]. Pour essayer de ne pas influencer la façon dont mes informateurs emploient la polydéfinitude, je ne leur ai pas dit que, au sein de l'ensemble de la langue, c'est cette structure en particulier que j'étudiais. La production langagière des locuteurs aurait sinon été biaisée donc non fiable (Skopeteas 2012 : 218).

Puisque la structure DEF N DEF A est optionnelle et peu fréquente, l'ensemble du corpus collecté n'a pas été transcrit, mais seulement les passages qui la comportaient. Le travail s'est effectué en cinq étapes :

1) écoute des fichiers audio en notant les passages comprenant DEF N DEF A
2) sélection et extraction des passages en question (une minute environ) avec Audacity
3) transcription des extraits
4) traduction en français
5) analyse comparative des syntagmes DEF N DEF A dans Libre Office Calc : type de déterminant, type d'adjectif, type de nom, place de l'adjectif, présence ou non d'une rupture entre les deux sous-parties du syntagme, répétition ou non de la préposition.

1.2. Les informateurs

Je me suis efforcée d'établir un équilibre entre informateurs femmes et hommes, jeunes, moins jeunes et anciens. Certains se sentaient plus à l'aise pour des entretiens concis, d'autres pour des narrations ou des récits de vie, d'autres encore dans des tâches plus ludiques comme les tests expérimentaux, d'autres pour des chants. J'ai donc panaché les types d'enregistrements selon les préférences des informateurs. Avec certains qui ne parlaient pas romani ou qui souhaitaient pratiquer des langues étrangères, j'ai parlé italien, anglais, serbe ou encore français.

Le profil des locuteurs ayant prononcé les exemples présentés dans ce chapitre fait l'objet du Tableau 154. Le Tableau 18 présente leur répartition par genre et par âge :

Tableau 18 : Répartition par genre et âge des locuteurs enregistrés dans le corpus spontané

	Femmes	**Hommes**	**-15 ans**	**15 – 39 ans**	**40 – 59 ans**	**+ 60 ans**
	9	9	0	8	5	5
Total	18		18			

	-15 ans		**15 – 39 ans**		**40 – 59 ans**		**+ 60 ans**	
	Femmes	**Hommes**	**Femmes**	**Hommes**	**Femmes**	**Hommes**	**Femmes**	**Hommes**
	0	0	4	4	2	3	3	2
Total	0		8		5		5	

[1] Il faudra vérifier, au moment de l'analyse, si ce facteur affecte l'émergence et l'usage de la structure qui nous intéresse.

Les enregistrements ont eu lieu dans l'environnement familier des locuteurs : à l'intérieur de la maison (salon, chambre isolée ou non), à l'extérieur (en promenade dans la rue, en terrasse de café) et en semi-extérieur (hall de maison, jardin, cour, dans la salle d'un café). Notons qu'ils ont rarement été réalisés de manière isolée : il était habituel que passent des membres de la famille ou amis, qui écoutaient l'entretien, voire le commentaient.

1.3. L'usage du romani au quotidien

1.3.1. Dans la famille

Mon hôte est né à Korçë en 1968. Il est journaliste, militant, délégué albanais auprès de l'Union Romani Internationale et vice-président du parlement de l'URI. L'arbre généalogique partiel de sa famille est construit par De Soto, Beddies et Gedeshi (2005 : 222), et reproduit en annexe 1, Figure 88 p. 540). J'ai pu recueillir les éléments biographiques suivants de ses ascendants :

- On sait peu de choses de son **arrière-grand père**, si ce n'est qu'il parlait le čergar II et était très riche et considéré parmi les Roms, si bien qu'il put acheter des terres.
- Son **grand-père**, de père et de mère čergars II, est né en 1890 à Kastoria (Grèce). Il a épousé une femme arli née en 1898 à Florina (Grèce), avant de « migrer »[1] à Korçë où ils eurent six enfants (deux filles et deux garçons, ainsi que deux qui n'ont pas survécu) nés entre 1920 et 1938. Il a, en parallèle, épousé une deuxième femme, elle aussi arli et née à Florina (Grèce) en 1900, avec qui il eut quatre enfants (une fille et trois garçons), tous nés également à Korçë entre 1924 et 1938.
- Sa **grand-mère** est la deuxième épouse : née à Florina, sa famille était originaire d'Izmir (Turquie) qu'elle avait fui suite à une épidémie. Selon la mémoire familiale, l'épidémie en question serait ensuite arrivée en Grèce, ce qui expliquerait que les familles auraient poursuivi la migration vers l'Albanie[2]. Elle aurait eu une première fille, née à Florina, avant son mariage.
- Les **enfants** de ce grand-père bigame ont tous parlé la variété de leur mère, c'est-à-dire l'arli, d'une part car ils ont été élevés principalement par leur mère et non par leur père, d'autre part en raison de l'environnement social. Les enfants grandissant dans un village ou un quartier strictement rom (*romani mahalla*), s'y sociabilisent dans la variété localement prédominante

[1] Il n'y avait pas, à l'époque, de frontières.

[2] Il s'agirait, selon mon hôte, d'une épidémie de choléra. Mais la synthèse établie par Hays (2005) montre que les vagues de choléra (quatrième pandémie : 1863-1875, cinquième : 1881-1896, sixième : 1899-1923) n'ont touché directement ni la Turquie ni la Grèce. D'autres fléaux sévissaient en revanche à l'époque. Zavitziano et al. (1900) fait état de diverses épidémies dans les « provinces turques », dont la scarlatine à Kavalla, la diphtérie, la grippe et la fièvre typhoïde à Constantinople en 1899-1900. Zavitziano (1901) rapporte qu'en avril 1901 a éclaté un début de peste bubonique à Constantinople. Ciccozzi et al. (2011) montrent enfin comment est née une épidémie d'hépatite C en Grèce au début de la décennie 1900 puis se serait propagée en Turquie et à Chypre jusque dans les années 1930.

– lorsqu'il y en a une dans le quartier. À Korçë, notamment dans le quartier rom et égyptien de Kullairi qui était le sien, vivaient une majorité d'Arlis, ce qui explique que la variété čergar II du père ait été moins ou moins bien transmise. Le grand-père de mon hôte était plus aisé, partant plus respecté que les autres. À l'époque du roi Zog, il servait d'intercesseur et de médiateur en cas de conflit avec la police et les autorités, ce pourquoi il portait le titre de Shaziman *Aga*[1]. Dans sa jeunesse, il se rendit trois fois aux États-Unis où il travailla comme ouvrier métallurgiste et comme maquignon. Il rencontra personnellement Enver Hoxha lorsque celui-ci vint visiter le quartier rom de Kullairi, et contribua à améliorer les conditions de vie du quartier rom en y faisant construire une école, un jardin, une route, aménager l'eau courante et un système d'égouts. Des immeubles collectifs et des maisons émergèrent au milieu des baraques.

- Le **père** de mon hôte, qui porte le nom de son propre grand-père, Demir, est né en 1928 à Korçë ; il y a épousé une femme rom, née en 1930 à Bilisht (Albanie), elle aussi du groupe arli. Ils eurent neuf enfants (six filles et trois garçons), tous nés à Korçë entre 1949 et 1970.

Mon hôte est le **huitième enfant** de ce père bivariétal arli/čergar II à variété très prédominante arli, et de cette mère arli de Bilisht. Il n'est locuteur que de la variété arli. Son père fut l'un des premiers Roms de la ville à emménager et à élever ses enfants dans un quartier non-rom, où sa maison était la seule maison rom – ce qui est un signe extérieur d'ascension sociale. Il eut, comme son père, beaucoup d'influence sur la communauté, servit de médiateur avec les autorités, arrangea les conflits, les mariages, etc. Il fut l'un de ces Roms qui purent continuer à faire des affaires et du profit privé même sous le régime.

L'**épouse** de mon hôte née à Elbasan en 1970. Elle parle la variété čergar I et l'arli, avec une prédominance du čergar I. Elle est femme au foyer et n'a jamais été à l'école. Elle a d'abord grandi à Elbasan avant de venir à Korçë à 14 ans, lorsqu'elle s'est mariée. Elle a deux sœurs et deux frères. Leur père, né à Elbasan et vivant aujourd'hui à Fier, est locuteur bivariétal du čergar I et de l'arli de par ses deux parents. Leur mère, née à Elbasan et ayant vécu à Peqin, était locutrice de la variété čergar I. Leur famille se portait très bien sous le régime, son grand-père possédait de nombreux chevaux. Comme sa mère est morte jeune, elle été élevée par ses grands-parents, son oncle ainsi que par la femme de celui-ci, qui est la sœur de son mari, locutrice de l'arli. Cela explique qu'elle ait été très tôt exposée à l'arli, avant même son mariage.

Leurs quatre enfants (trois filles et un garçon) sont aujourd'hui bivariétaux, avec pour variété prédominante l'arli, celle de leur père et celle de la majorité des Roms de Korçë. Ils ont également grandi en ville et dans deux maisons successives de quartiers à prédominance non-rom.

[1] *Aga*, du turc ottoman *aga*, était un titre honorifique des officiers et hauts fonctionnaires dans l'Empire ottoman. Dans les provinces, c'est donc l'équivalent de « seigneur » pour désigner un notable ou un dignitaire, notamment après la chute de l'empire.

1.3.2. Hors de la famille

Si la famille qui m'hébergeait est bilingue romani et albanais, j'ai pu constater que ce n'était pas le cas de toutes. La variation individuelle est bien sûr très importante voire déterminante, mais l'usage spontané de la langue et la « confiance en soi » linguistique m'ont aussi semblé aller globalement croissants selon les facteurs : de l'âge (plutôt les plus de 25 ans, et surtout les plus de 40 ans), du genre (plutôt les femmes), du prestige social et de la richesse (plutôt les personnes fortunées ou jouissant de responsabilités), du groupe (plutôt les Arlis et les Čergars I), du militantisme linguistique ou professionnel, et bien sûr du contenu de la discussion et du contexte d'énonciation (lieu, personnes présentes...). Certains événements relevant des domaine de la famille et de la communauté (mariage, naissance, décès et tous les préparatifs afférents), l'éducation des jeunes enfants ou le soin des morts au cimetière, sont exclusivement menés en romani. Ces observations personnelles demanderaient à être complétées par une étude ethno- et sociolinguistique approfondie.

Du fait des inter-mariages et/ou des relations professionnelles, un certain nombre des personnes connaissaient plusieurs variantes de la langue : arli et čergar I, arli et bam(b)il, čergar I et mečkar... J'ai au contraire rencontré des personnes (notamment mečkars, et quelques Arlis) estimant que la langue ne leur avait pas été suffisamment transmise dans le cadre familial, et qui ont suivi des cours de langue auprès d'associations. Ils correspondent en partie au concept de néo-locuteurs, *new speakers*, forgé entre autres par O'Rourke et Ramallo (2013). Nous avons soulevé en section 2.4.6, p. 42, les points communs et les différences qui existent avec la situation des néo-locuteurs d'Irlande ou de Galicie. Les néo-locuteurs albanais qui apprennent ou réapprennent le romani dans un tel cadre parlent, pour certains, une langue mâtinée de plusieurs variétés, que l'on peut considérer comme une variété à part entière, qualifiée par elles-mêmes de « standard »[1]. Cette transmission différente de la langue, qui s'effectue par un apprentissage délibéré et un cadre institutionnel, a-t-elle une influence sur l'usage de la structure polydéfinie ? Une telle *koiné* connaît-elle cette structure ? La question se pose également pour les militants de l'élite rom qui ont une connaissance large des différents variétés d'Albanie et/ou des autres pays, et en ce qui concerne les enseignants de romani, qui pourraient éventuellement adopter une approche normative de la langue. J'ai par exemple noté que des militants profitaient de l'occasion offerte par le test de localisation pour dispenser une petite « leçon » de vocabulaire ou de dialectologie à l'autre participant (et/ou à moi-même). Parfois ils souhaitaient rechercher le nom des objets dans un dictionnaire[2] plutôt que d'employer le mot qui leur venait spontanément à l'esprit, si celui-ci était albanais. J'ai choisi de ne pas écarter ces locuteurs du corpus, car ils font partie du paysage linguistique local.

[1] Le terme *standard* était spontanément employé par nos informateurs pour désigner soit une variété considérée comme prestigieuse (par exemple l'arli ou le čergar), soit la *koiné* pan-européenne parlée par une partie de l'élite militante. Cette dernière présente une forme fluide et très différente d'un pays et d'un locuteur à l'autre.

[2] Le dictionnaire dont disposent certains militants, le seul disponible en albanais, est Courthiade (2004), Courthiade (2009b) n'étant pas encore beaucoup en circulation à l'époque.

1.4. Le matériel et les données collectées

Le corpus spontané a été enregistré soit de manière formelle avec un enregistreur Tascam et le micro serre-tête et/ou le micro statique, soit de manière informelle avec un petit dictaphone Olympus portatif, léger, discret et sans fil. Tous deux étaient posés sur la table lorsqu'il y en avait une ou bien tenus à la main lorsqu'il n'y en avait pas ou que le bruit de fond était important. Plusieurs types d'enregistrement ont été réalisés :

- conversations spontanées (en ma présence ou non)
- entretiens dirigés ou semi-dirigés permettant notamment de collecter des méta-données. Ils ont été menés soit par moi soit par mes collaborateurs, en ma présence ou non,
- narrations : récits de vie, anecdotes, contes,
- textes procéduraux : recettes de cuisine,
- chansons.

Les enregistrement ont eu lieu chez les Roms concernés, dans la maison ou dans la cour, ou bien encore à l'extérieur, au café ou au parc.

Le Tableau 19 présente la composition du corpus spontané. Une partie du corpus n'a pas été enregistrée car elle est issue d'observations participantes, ou encore de conversations du quotidien[1] – ce qui est l'avantage et l'inconvénient d'un terrain en immersion dans une famille.

[1] Je n'avais pas toujours le microphone sous la main ni allumé pendant que l'on prenait nos repas, que je faisais ma toilette ou les courses.

Tableau 19 : Nombre d'enregistrements effectués en Albanie

		Korçë	Voskopojë	Tiranë	Bilisht	Ohrid	Llakatund	Berat	Fier	Tushemisht	Total
Conversation spontanée	**2013**	55	11	19	-	-	-	-	-	-	**85**
	2014	52	2	-	2	11	8	25	-	6	**106**
Entretiens et récits de vie	**2013**	25	12	11	-	-	-	-	-	-	**28**
	2014	43	5	-	12	3	17	23	-	10	**113**
Chants	**2013**	28	21	9	-	-	-	-	-	-	**58**
	2014	13	-	-	-	-	2	9	1	-	**25**
Contes	**2014**	10	-	-	-	-	-	-	-	1	**11**
Éléments de vocabulaire isolés	**2013**	2	4	1	-	-	-	-	-	-	**7**
	2014	24	-	1	7	9	24	5	3	-	**73**
Recettes de cuisine	**2014**	10	-	-	1	-	-	-	-	1	**12**
Élicitation sur la polydéfinitude ou enquête dialectologique	**2013**	2	-	-	-	-	-	-	-	-	**2**
	2014	6	-	-	-	1	-	-	-	-	**7**
Cours de langue	**2013**	-	2	2	-	-	-	-	-	-	**4**
Test d'instruction et de localisation	**2014**	53	8	-	4	3	10	2	-	-	**80**
Discussion précédant le test	**2014**	18	2	-	-	2	1	-	-	-	**23**

1.5. La transcription et la présentation des exemples

Les données spontanées et semi-spontanées ont commencé à être transcrites sous Libre Office Writer sur place également et suivant l'orthographe albanaise. Une cinquantaine d'exemples issus du corpus spontané sera présentée au cours de l'analyse. Seules les phrases d'exemples sont ici citées, avec glose et traduction ; le corpus lui-même, transcrit et traduit, se trouve en annexe (sections 1.2.4, p. 549, à 1.2.11). Les exemples du corpus spontané comme du corpus semi-spontané sont présentés de la façon suivante :

(n°) *Locuteur, Ville, Date de l'enregistrement*

La	*phrase*	*en*	*romani.*
GLOSE	GLOSE	GLOSE	GLOSE

« Traduction en français. »

Sur la première ligne sont indiqués, de gauche à droite : le numéro de l'exemple, le ou la locuteur qui prononce la phrase, le lieu et date de l'enregistrement, et enfin la page où l'on peut retrouver le corpus. Cette dernière est un hyperlien qui renvoie au texte intégral situé en annexe, c'est-à-dire l'extrait de conversation transcrit en romani, traduit en français et analysé. Les hyperliens permettent un va et vient entre la partie principale du présent travail et les annexes.

La deuxième ligne est écrite en romani, suivant les conventions orthographiques des linguistes du romani[1]. La troisième ligne consiste en une glose morpho-syntaxique interlinéaire, effectuée morphème par morphème suivant les conventions de Leipzig dans sa version révisée de 2008[2] puis de 2015[3]. Les abréviations sont données en français (cf. Table des abréviations). La quatrième ligne est une traduction en français. Certaines syllabes du discours sont inaudibles à l'écoute postérieure de l'enregistrement : je les ai remplacées par un X dans le texte – un X pour une syllabe. Lorsqu'elles sont audibles mais que je n'ai pas réussi à les traduire, même avec l'aide de mes informateurs, je les ai transcrites mais glosées par un point d'interrogation.

Les locuteurs sont anonymisés de la façon suivante :

- une lettre représentant le genre féminin (F) ou masculin (M) de la personne
- un nombre représentant l'âge numérique du locuteur au moment de l'enregistrement

[1] Lorsque l'accent du locuteur est très marqué, la transcription peut comprendre des symboles issus de l'Alphabet Phonétique International (ə, ɽ) dans sa version révisée de 2015 : https://upload.wikimedia.org/wikipedia/commons/e/e4/The_International_Phonetic_Alphabet_%28revised_to_2015%29.pdf

[2] http://www.eva.mpg.de/lingua/pdf/LGR08.02.05.pdf

[3] https://www.eva.mpg.de/lingua/pdf/Glossing-Rules.pdf

- une lettre représentant la variété albanaise parlée par la personne : a (néo-locuteur arli), M (mečkar), A (arli), Č (čergar I), AČ (arli & čergar I / čergar I & arli), AB (arli & bamlija), MČ (mečkar & čergar I), MB (mečkar & bamlija).

Ainsi « F15A » code-t-il une femme de 15 ans parlant la variété arli. L'abréviation « Ch » code *chercheuse* et introduit les énoncés que j'ai moi-même produits. Les hésitations, pauses et reprises sont représentées par une barre oblique /.

1.6. Analyse des données

Le Tableau 20 présente la typologie que l'on suivra présentée pour le statut de l'information (comprenant trois grands types de statut : donné, accessible et nouveau), ainsi celle de la structure de l'information (comprenant deux grands types de topique et deux grands types de focus). Nous distinguons donc statut et structure de l'information, comme Götze et al. (2007 : 150), mais aussi comme Szendrői (2010 : 869).

Tableau 20 : *Tags of the Extended Annotation Scheme for Information Structure* (Götze et al. 2007 : 148-150)

Layer	Tag	Short description
Information Status	giv	given (underspecified)
	giv-active	active
	giv-inactive	inactive
	acc	accessible (underspecified)
	acc-sit	situationally accessible
	acc-aggr	aggregation
	acc-inf	inferable
	acc-gen	general
	new	new
Topic	ab-t	aboutness topic: what the sentence is about
	fs-t	frame-setting topic: frame within which the main predication holds
Focus	nf	new-information focus (underspecified)
	nf-sol	solicited new-information focus
	nf-unsol	unsolicited new-information focus
	cf	contrastive focus (underspecified)
	cf-repl	replacement - contrastive focus
	cf-sel	selection - contrastive focus
	cf-part	partiality - contrastive focus
	cf-impl	implication - contrastive focus
	cf-ver	truth value (verum)- contrastive focus
	...+op	All kinds of foci given above can occur as bound by focus operators like the particles only, even, also etc. as well as negation operators.

A partir d'exemples issus de discours spontané, nous nous efforcerons de ne pas séparer, dans l'analyse, ce qui ressort de la syntaxe, de la sémantique, et des fonctions de la structure DEF N DEF A. Nous sommes à la recherche d'un cadre théorique général, qui puisse notamment expliquer le caractère optionnel de la polydéfinitude. Les locuteurs peuvent en effet construire de huit manières le syntagme comprenant les éléments {article défini ; nom ; modifieur adjectival} pour exprimer par exemple « le grand frère » (Tableau 21).

Tableau 21 : Huit structurations possibles du syntagme nominal avec adjectif épithète

	A antéposé	**A postposé**
un ART.DEF	DEF A N *o baro phral*	DEF N A *o phral baro*
plusieurs ART.DEF	DEF A DEF N *o baro o phral*	DEF N DEF A *o phral o baro*
un ART.DEF + rupture	DEF A/ N *o baro, phral*	DEF N/ A *o phral, baro*
plusieurs ART.DEF + rupture	DEF A/ DEF N *o baro, o phral*	DEF N/ DEF A *o phral, o baro*

La barre oblique représente ici une rupture : une pause ou l'insertion d'autres mots au sein du syntagme. Cette rupture doit se situer à la frontière entre la partie « nominale » et la partie « adjectivale ». Nous avons proposé ici des structures où le premier déterminant est un article défini. Il peut néanmoins être un autre déterminant, possessif ou démonstratif.

Nous pensons, avec Nølke (2001 : 110), que « le locuteur s'efforce toujours de faire passer son message dans une forme qui corresponde à ses intentions, et pour ce faire, il a une gamme de moyens linguistiques à sa disposition ». Les structures linguistiques servent un but communicatif : il faut donc chercher l'explication dans la valeur communicative des structures individuelles (Nølke 2001 : 172). Qu'est-ce qui motive le choix des locuteurs, lorsqu'ils choisissent dans la chaîne de parole une construction parmi les huit possibles du Tableau 21 ?

Étant donné qu'il existe en romani des syntagmes DEF N A (postposition de l'adjectif sans répétition de l'article) et qu'il existe extrêmement peu, sinon pas du tout, de syntagmes DEF A DEF N (répétition de l'article sans postposition de l'adjectif), tout laisse à croire que la polydéfinitude (DEF N DEF A) relève de deux phénomènes distincts mais combinés : la postposition de l'adjectif et la duplication de l'article défini. Il nous faut comprendre si la postposition de l'adjectif est inéluctable et va donc nécessairement de pair avec la duplication de l'article défini. Il nous faut également comprendre ce qui, du point de vue pragmatique, motive cette duplication : pourquoi la postposition de l'adjectif ne « suffit »-elle pas ? Ces syntagmes présentent des nuances et effets de sens différents : ils ne sont donc pas destinés à apparaître dans les mêmes contextes. Le présent chapitre se donne pour objet de déterminer quels sont ces contextes.

2. Syntagmes sans rupture interne

2.1. Syntagmes polydéfinis

Le fil rouge de cette section est le focus (Campos et Stavrou 2004 : 141-142) :

> « *Crucially, polydefinites imply a contrast and are organized around the presupposition/focus distinction.* […] *In a polydefinite construction, the article-noun sequence constitutes the context or presupposition and the articled adjective constitutes the focus (or the 'assertion,' in Zubizarreta's 1998 terms). In other words, the definite article of the adjective in a polydefinite construction undertakes the same function as contrastive focus on the adjective in a monadic DP, the overall effect being one of contrast or selection among the referents of the denotation of the noun which, by not being stressed, constitutes the context or a contextually retrievable set*. »

Pour la polydéfinitude en romani, Boretzky (1994 : 55) parlait en effet d'emphase et Boretzky (2000 : 42-43) de focus sur l'adjectif.

2.1.1. DEF N DEF A

Notre corpus révèle trois types de syntagmes DEF N DEF A : ceux qui conferment l'assertion « *the overall effect being one of contrast or selection among the referents of the denotation of the noun* », ceux qui produisent un autre effet, que nous nommerons pour l'instant d'emphase, et ceux dont l'interprétation est ambiguë entre effet de sélection et d'emphase. Le corpus transcrit et traduit, ainsi que le détail de l'analyse de chaque exemple, se trouve en section 1.2.4., p. 549.

2.1.1.1. Effet de sélection

Dans les exemples (154) à (161), l'adjectif revêt un rôle de sélection. Nous développons l'exemple (154) avec son corpus, sa glose et son analyse détaillée. Pour plus de fluidité, nous ne présenterons ensuite que la glose des exemples suivie de la synthèse des analyses – le corpus comme les analyses détaillées se trouvant en annexe et faisant l'objet d'un renvoi. Si un point particulier nécessite un éclaircissement, il se trouve à la fois dans le corps de l'étude et dans l'analyse détaillée en annexe.

L'exemple (154) est issu d'un entretien dirigé, où la locutrice présente sa famille et énumère ses frères et sœurs.

1	*F60A: Me daja čardena la Pakize.*
2	*M46A: Keći maksemja isi te daja?*

3	*F60A: Me daja isi la enja maksemja.*
4	*M46A: Ka vaćeres mang lenđe lafja/ sare maksemenđe.*
5	*F60A: Mi/ mi bari phen isi Demirana.*
6	*M46A: Hmhm.*
7	*F60A: O/ odja vaver čardena la Fidarija.*
8	*E/ odəja i trin kaj si mi phen čardena la Nife.*
9	*Me isinom i štar/*
10	*M46A: štarenđiri.*
11	*F60A: E štarengi isinom me Drita čardena man.*
12	*Sora isi man daa jek vaver phen/ čardena la: Moza.*
13	*M46A: Hmhm.*
14	*F60A: Daa jek i tikni (i?) xurdi/ isi: Anida/ čardena la.*
15	*Sora isi mo baro phral/ kaj čardena les Yli.*
16	*M46A: Hmhm.*
17	*Isi o phral o vaver/ pala leste/ čardena les Šaziman thodam amare papusko lafi lesće.*
18	*M46A: šukar.*
19	*F60A: Ah/ isi man daa jek vaver phral/ čardena les Arben.*
20	*M46A: Te oves sasto baxtalo.*
1	F60A : Ma mère s'appelle Pakize.
2	M46A : Combien d'enfants a ta mère ?
3	F60A : Ma mère elle a neuf enfants.
4	M46A : Tu vas me dire leur nom/ à tous les enfants.
5	F60A : Ma/ ma grande soeur est Demirana.
6	M46A : Hmhm.
7	F60A : le/ l'autre s'appelle Fidaria de son prénom.
8	Euh/ la troisième qui est ma soeur elle s'appelle Nifa.
9	Moi je suis la quatre/
10	M46A : La quatrième.
11	F60A : La quatrième c'est moi, je m'appelle Drita.
12	Ensuite j'ai encore une autre sœur/ elle s'appelle : Moza.
13	M46A : Hmhm.
14	F60A : Encore une la cadette la benjamine/ c'est : Anida elle s'appelle.
15	Ensuite il y a mon grand frère/ qui s'appelle Yli.
16	M46A : Hmhm.
17	F60A : Il y a l'autre frère/ après lui/ il s'appelle Shaziman/ on lui a donné le nom de notre grand-père.
18	M46A : Bien.
19	F60A : J'ai encore un autre frère/ il s'appelle Arben.
20	M46A : Merci à toi.

(154) *F60A, Korçë, 21 juillet 2014*

Isi	***o***	***phral***	***o***	***vaver,***	*pala*	*les-te,*
être.3.PRES	ART.DEF.M.SG	frère(M.SG)	ART.DEF.M.SG	autre(M.SG)	après	3.SG.OBL-LOC
čard-ena	*les*	*Šaziman.*				
appeler-3.PL.PRES	3.SG.OBL	Shaziman				

« Il y a un autre frère après lui, il s'appelle Shaziman. »

Statut de l'information

Le référent « l'autre frère » est accessible-inférable (*set relation*) puisque la locutrice a annoncé qu'elle avait huit frères et sœurs. L'adjectif *autre* est nouveau lexicalement et le nom *frère* donné-actif lexicalement.

Structure de l'information

Le syntagme *o phral o vaver*, l'autre frère, constitue le focus de la phrase, plus précisément un focus contrastif du type *subtype selection* chez Götze et al. (2007 : 180-181).

Interprétation sémantique

La partie nominale du syntagme, le frère, fait partie du contexte, du présupposé. La locutrice a en effet annoncé qu'elle avait de nombreux frères et sœurs et vient, dans la phrase qui précède, de mentionner le plus grand de ses frères (qui est en fait plus jeune qu'elle). Il s'agit ensuite de sélectionner un autre frère. Le contraste se fait donc *in presentia* en discours. Cette sélection se fait au moyen de l'adjectif, qui apporte un sème spécifique permettant d'identifier le référent. La référence se calcule par l'adjonction des sèmes de l'adjectif et de ceux du nom. En d'autres termes, le référent de l'expression globale est à l'intersection de l'ensemble {frères} et de l'ensemble {autres entités}. L'identification du référent se fait en opérant une restriction pragmatique, au sein des référents du nom grâce à l'adjectif.

(155) *M18Č, Korçë, 24 juillet 2014*

M-o	***phral***	***o***	***xurd-o***	*džan-el*	*angliśt.*
POSS.1SG-M.SG	frère(M.SG)	ART.DEF.M.SG	tout_petit-M.SG	savoir-3SG.PRES	en_anglais

« Mon petit frère parle anglais. »

(156) *M24M, Llakatund, 12 août 2014*

Se/	***m-i***	***phen***	***i/***	***i***	***tikn-i,***	*beš-ela*	*Germanija*
parce que	POSS.1SG-F.SG	frère(F.SG)	ART.DEF.F.SG	ART.DEF.F.SG	petit-F.SG	habiter-3SG.PRES	Allemagne

« Et/ ma/ ma petite sœur, habite en Allemagne. »

Notons que pour l'exemple (156), le syntagme ne présente pas de rupture interne car l'hésitation a lieu non entre la partie « nominale » et la partie « adjectivale » du syntagme polydéfini, mais au sein de la partie adjectivale. Le locuteur hésite non sur le fait d'employer la polydéfinitude, mais sur la propriété distinctive à mentionner.

(157) *M24M, Llakatund, 12 août 2014*

ə/	*ćer-ena/*	*pandž*	*breš*	***m-o***	***phral***	***o***	***tikn-o.***
euh	faire-3PL.PRES	cinq	année[M.PL]	POSS.1SG-M.SG	frère(M.SG)	ART.DEF.M.SG	petit-M.SG

« Euh/ ça fait/ cinq ans/ mon petit frère. »

Dans l'exemple (157), l'identification du référent ne se fait pas en opérant une restriction pragmatique au sein des référents du nom, à l'intersection de l'ensemble {frères} et de l'ensemble {entités petites}, car le locuteur n'a qu'un seul frère. L'adjectif n'effectue pas à proprement parler une restriction au sein de l'ensemble des frères. *Petit*, qui constitue le focus, n'est donc pas une propriété permettant de contraster le frère avec un autre frère, c'est-à-dire permettant l'identification. L'adjectif transmet une information certes, (le fait que le frère soit plus jeune que le locuteur), mais une information non-restrictive. C'est le cas seulement si l'identification consiste à sélectionner le frère parmi l'ensemble {frères et sœurs}, c'est-à-dire un ensemble plus large que celui du nom ou de l'adjectif. L'adjectif effectue donc une sélection, non au sein de l'ensemble des frères masculins, mais au sein de l'ensemble des {frères et sœurs}. La Figure 30 représente cette opération de sélection.

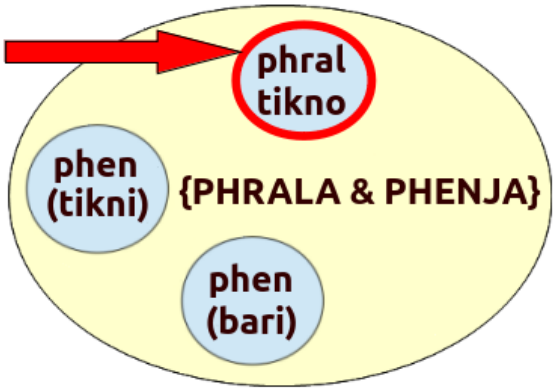


Figure 30 : Sélection d'un petit frère parmi les frères et sœurs

(158) *M71A, Korçë, 28 juillet 2014*

Akova	*o*	*bar-o*	*raklo/*	***o/***
DEM.M.SG	ART.DEF.M.SG	grand-M.SG	garcon_non_rom(M.SG)	ART.DEF.M.SG
o	***phral***	***o***	***bar-o/***	*mukh-l-as*
ART.DEF.M.SG	frère(M.SG)	ART.DEF.M.SG	grand-M.SG	laisser-PRET-3.SG
p-e	*romn-ja*	*deodžende/*	*khamn-i/*	*džan-esa.*
POSS.REFL-F.SG.OBL	femme-F.SG.OBL	?	enceint-F.SG	savoir-2SG.PRES

« Ce grand garçon, le… le frère aîné laissa sa femme (?)/ enceinte, tu sais. »

L'exemple (158) nous montre que la structure polydéfinie peut avoir une fonction supplémentaire. Comment expliquer que le référent, qui était jusque là désigné comme *o baro phral*, DEF A N, soit cette fois désigné comme *o phral o baro*, DEF N DEF A ? C'est qu'entre temps, il a reçu une autre

désignation encore, *o baro raklo*, le grand garçon non-rom. On apprend donc que le personnage en question n'est pas Rom, mais l'on perd le lien de parenté avec les autres frères. Peut-être de peur que j'identifie mal le personnage (et que je le confonde par exemple avec l'homme qui vend de la pensée), il s'auto-corrige en juxtaposant un deuxième syntagme. Pour clarifier les choses, il réintroduit alors le terme de *frère*, cette fois en introduisant un focus étroit sur l'adjectif, d'où la postposition de l'adjectif : *o phral o baro*, le grand frère. La structure DEF N DEF A, en restreignant la référence à une entité au sein d'un ensemble, peut servir à préciser ce qui est dit, à apporter une correction.

Cela explique selon nous pourquoi l'adjectif est postposé, alors même qu'il est donné actif. Pour apporter sa clarification, le locuteur doit mettre l'adjectif en position focale, en faisant semblant que la propriété de l'adjectif est nouvelle, alors qu'elle est en réalité déjà connue. On voit qu'il n'y a pas d'incompatibilité entre le focus et le statut donné de l'information : il suffit que le locuteur choisisse de ne plus présenter comme donnée ou connue une information pour que la focalisation devienne possible, même s'il s'agit en réalité d'un savoir partagé. Pour ce faire, il a à sa disposition la structure dite polydéfinie. Nølke (2001 : 112) décrit ce type de stratégie :

> « Le choix entre la présentation de l'information (en principe) factuelle et connue comme vieille ou comme nouvelle n'est pas complètement libre. Il dépend du degré de proéminence cognitive de l'information. Si le locuteur désire répéter quelque chose qui vient d'être mentionné, il lui faudra le présenter comme de l'information connue, mais il pourra le focaliser. En effet, s'il le présente comme de l'information focalisée *nouvelle*, la réaction de l'allocutaire pourra être quelque chose comme : « mais tu sais bien que je le sais déjà ! ». Cet artifice pourra évidemment servir à provoquer des inférences, car selon le principe de pertinence, l'allocutaire réagira probablement en pensant : « Pourquoi me raconte-t-il quelque chose qu'il sait que je sais ? » Nølke (2001 : 112)

(159) *F38A, Korçë, 4 août 2014*

Akəna	*al-i*	*i*	*čhaj*	*kat-o*	*kampi/*
maintenant	venir.PRET-3SG.F	ART.DEF.F.SG	fille(M.SG)	venant_de-ART.DEF.M.SG	camp[1](M.SG)
i	***čhaj***	***i***	***vaver***.		
ART.DEF.F.SG	fille(M.SG)	ART.DEF.F.SG	autre(F.SG)		

« La fille vient de rentrer du camp / l'autre fille. »

(160) *M23M, Llakatund, 13 août 2014*

Važdon-av	*sa*	*opre/*	*i/*	*t-**i***	***škola***	***i***	***bar-i***.
terminer-1SG.PRES	tout	haut	ART.DEF.F.SG	à-ART.DEF.F.SG	école(F.SG)	ART.DEF.F.SG	haut -F.SG

« Je la fais jusqu'au bout/ la/ la grande école. »

Dans l'exemple (160), l'adjectif tend à se lexicaliser avec le nom pour former un quasi mot composé.

[1] Activités de type centre aéré, organisées gratuitement par l'OSCE pendant quelques jours au centre de la ville, à destination des enfants les plus pauvres de la ville et de sa périphérie.

(161) *M46A, Korçë, 24 juillet 2014*

O	***phral***	***o***	***tikn-o/***
ART.DEF.M.SG	frère(M.SG)	ART.DEF.M.SG	petit-M.SG
beš-ela	*isine/*	*e*	*guruv-en-sar.*
habiter-3SG.PRES	être.3.PRET	ART.DEF.PL.OBL	vache- M.PL.OBL-INSTR

« Le petit frère/ habitait/ avec les vaches. »

Dans cet exemple, on peut supposer qu'il existe un focus étroit de contraste entre ce référent et les deux autres frères : l'un est petit, les autres grands ; l'un travaille à la maison, les autres à l'étranger ; l'un travaille avec ses vaches, les autres avec et pour un paysan non-rom. Mais ce focus étroit semble avoir une autre fonction, celle d'attirer l'attention sur un changement de point de vue, au sein du récit. En effet le benjamin n'est pas encore entré en action dans l'histoire, contrairement aux deux frères aînés, qui travaillent à l'étranger pour un Non-Rom et gagnent de l'argent. Le centre d'attention du récit se porte ensuite sur le benjamin, qui n'est plus désigné comme le fou mais simplement comme le petit frère. Il ne sera plus question, par la suite, des deux frères aînés. La polydéfinitude permet d'attirer l'attention de l'interlocuteur sur le *topic-shift*, le changement de sujet et/ou d'objet dans le cours du récit.

L'analyse informationnelle et sémantique donne les résultats présentés dans le Tableau 22. Si le référent de l'expression polydéfinie ainsi que le nom lexical, sont donnés ou accessibles, l'adjectif est souvent nouveau du point de vue lexical. Le syntagme constitue le topique de la phrase. L'interprétation de l'expression se fait à l'intersection des ensembles de référents, grâce à une restriction pragmatique. On a pu observer deux autres fonctions de la polydéfinitude : un artifice cognitif pour rendre proéminente une information qui est connue, mais que le locuteur veut réactiver, et le *topic-shift*. Le contraste se fait toujours *in presentia*.

Tableau 22 : Analyse informationnelle et sémantique des syntagmes DEF N DEF A à effet de sélection

DEF N DEF A	Statut de l'information			Structure de l'information	Interprétation sémantique	Contraste
	référent du syntagme	**nom (lexicalement)**	**adjectif (lexicalement)**			
(154)	accessible-inférable (*set relation*)	donné	nouveau	focus contrastif (*selection)*	intersection restriction pragmatique présupposé (N) sélection et focus étroit (A)	*in presentia*
(155)	accessible-inférable (*set relation*)	accessible-inférable	nouveau	topique-propos	intersection restriction pragmatique présupposé (N) sélection et focus étroit (A)	*in presentia*
(156)	accessible-inférable (*set relation*)	donné-inactif	nouveau	topique-propos	intersection restriction pragmatique présupposé (N) sélection et focus étroit (A)	*in presentia*
(157)	donné-inactif	donné-inactif	donné	topique-propos	pas d'intersection ni de restriction pragmatique, sauf si l'ensemble est plus large que le seul N focus étroit (A)	*in presentia*
(158)	donné-actif	donné-inactif	donné	topique-propos	intersection restriction pragmatique focus étroit (A) correction artifice cognitif	*in presentia*
(159)	donné-actif	donné	nouveau	topique-propos	intersection restriction pragmatique présupposé (N) focus étroit (A)	*in presentia*
(160)	donné-actif	donné	nouveau	topique-cadre	intersection ? restriction pragmatique présupposé (N) expression ambiguë	*in presentia*
(161)	donné-inactif	donné	donné-inactif	topique-propos	intersection restriction pragmatique présupposé (N) focus étroit (A) *topic-shift*	*in presentia*

2.1.1.2. Effet d'emphase

Nous présentons ici deux exemples où l'adjectif revêt un rôle d'emphase.

(162) *F16AČ, Korçë, 21 juillet 2014*

Istanbul/	*kaj*	*ćer-ena*	***o***	***film-e***	***o***	***bar-e!***
Istamboul	où	faire-3PL.PRES	ART.DEF.PL	film-PL	ART.DEF.PL	grand-PL

« Istamboul, là où on fait les grands films ! »

(163) *M46A, Korçë, 15 août 2014*

Aćija	*si*	***i***	***bira***	***i***	***puran-i.***
DEM.F.SG	être.3.PRES	ART.DEF.F.SG	brasserie(F.SG)	ART.DEF.F.SG	vieux-F.SG

« Ça c'est l'ancienne brasserie. »

Figure 31 : La brasserie de Korçë

Dans l'exemple (163), l'adjectif « ancienne » n'est pas restrictif car la ville de Korçë n'a jamais eu qu'une seule fabrique de bière. Il donne une propriété du référent.

L'analyse informationnelle et sémantique donne les résultats suivants (Tableau 23) :

Tableau 23 : Analyse informationnelle et sémantique des syntagmes DEF N DEF A à effet d'emphase

DEF N DEF A	Statut de l'information			Structure de l'information	Interprétation sémantique	Contraste
	référent du syntagme	nom (lexicalement)	adjectif (lexicalement)			
(162)	nouveau ou accessible-général	nouveau	nouveau	focus de nouvelle information non sollicitée	focus étroit assertif emphase	*in absentia*
(163)	accessible-situationnel	nouveau	nouveau	focus de nouvelle information non sollicitée	focus assertif	*in absentia*

Si le référent de l'expression polydéfinie semble plutôt accessible, les termes employés sont en revanche nouveaux du point de vue lexical. Le syntagme constitue le focus de la phrase. L'interprétation de l'expression est celle d'un focus assertif, qui va de pair avec le fait que la phrase soit exclamative, dans l'exemple (162), et que le référent soit annoncé par un déictique en début d'énoncé et physiquement montré du doigt, dans l'exemple (163). Le contraste se fait toujours *in absentia*.

2.1.1.3. Effet hybride

Les exemples (164) à (168) peuvent s'interpréter de deux manières différentes. À chaque fois, une interprétation avec effet de sélection et une interprétation à effet d'emphase sont possibles. Nous pensons qu'elles ne s'excluent pas l'une l'autre, mais qu'elles coexistent. Une expression peut avoir plusieurs visées communicatives concomitantes : le contraste et la focalisation interprétative.

(164) *F16AČ, Korçë, 24 juillet 2014*

Me	*bian-dil-om*	*ando*	*ćher*	*odothe,*	*kaj*
1.SG	accoucher-PASS.PRET-1SG	dans.ART.DEF.M.SG	maison(M.SG)	là-bas	où
si	*akana*	***o***	***lokali***	***o***	***bar-o***.
être.3.PRES	maintenant	ART.DEF.M.SG	café(M.SG)	ART.DEF.M.SG	grand-M.SG

« Je suis née dans la maison là-bas, où il y a maintenant le grand café. »

(165) *F16AČ, Korçë, 24 juillet 2014*

Kaj	*isi*	***o***	***lokal***	***o***	***bar-o/***	*ti/*
où	être.3.PRES	ART.DEF.M.SG	café(M.SG)	ART.DEF.M.SG	grand-M.SG	à.ART.DEF.F.SG
(166) *katredale.*	*Othe*	*isine*	***amar-o***	***ćher***	***o***	***bar-o***.
cathédrale(F.SG)	Là	être.3.PRET	POSS.1PL-M.SG	café(M.SG)	ART.DEF.M.SG	grand-M.SG

« Là où il y a le grand café/ à la/ cathédrale. Là il y avait notre grande maison »

Pour ces occurrences, on peut penser que l'effet d'emphase prédomine sur l'effet de sélection. La locutrice me signifie implicitement que la maison de sa prime enfance lui plaît mieux que celle dans

laquelle elle a déménagé ensuite. Elle met donc sous emphase la propriété 'grand' de l'ancien référent, qui prend un sens de superlatif absolu (notre *très grande* maison) et non relatif (notre maison *la plus grande [des deux]*). La maison dans laquelle elle vit actuellement m'a même semblé plus spacieuse que celle de sa prime enfance. Il est selon nous important de prendre en compte la subjectivité de la locutrice.

La répétition de l'adjectif *baro*, grand, pour désigner à la fois la maison et le café, donne l'impression que la propriété d'être grand est passée de la maison d'enfance au café qui occupe à présent le rez-de-chaussée de l'immeuble. Il semble y avoir transmission de cette propriété de grandeur : si tel est le cas, il s'agit d'une propriété intrinsèque du référent désignée, en dehors de toute comparaison avec les cafés voisins dans la rue.

On peut dire que la référence se construit au fur et à mesure des phrases, puisque trois syntagmes nominaux sont successivement utilisés pour désigner le même référent. Si l'on fait abstraction de la pause et du bâillement, la locutrice prononce les phrases suivantes :

Akate isi amaro ćher/	*kaj isi o lokal o baro/ ti/ katredale.*	*Othe isine amaro ćher o baro.*
POSS N	DEF N DEF A	POSS N DEF A
Ici c'est notre maison/ là où il y a le grand café/ à la/ cathédrale. Là il y avait notre grande maison.		

Figure 32 : SN polydéfini synthétisant deux SN précédents

La Figure 32 montre que le troisième syntagme représente comme une fusion des deux précédents, du point de vue grammatical.

(167) *F66M, Voskopojë, 28 juillet 2014*

Ose	*nane*	*problemi*	*o*	*romnj-a-k-o*	
ou	être.NEG.3.PRES	problème	ART.DEF.M.SG	femme_Rom-OBL.SG-GEN-M.SG	
keq/	*po*	*but*	*isi*	*o*	*rom-a*
mal(M.SG)	plus	beaucoup	être.3.PRES	ART.DEF.PL	homme_Rom-M.PL
tha	***o***	***čhav-e***	***o***	***bar-e.***	
et	ART.DEF.PL	enfant-M.PL	ART.DEF.PL	grand-PL	

« Et le problème ne vient pas des femmes/ mais bien plutôt des hommes et des grands enfants […]. »

(168) *M68A, Voskopojë, 15 juillet 2013*

Isin-am	*Rom-a/*	*kaj/*	*dž-asa/*	*pal/*
être-1PL.PRES	Rom-M.PL	REL	aller-1PL	après
t-o	***ćidipe***	***t-o/***	***gadžikan-e.***	
à.ART.DEF.PL	rassemblement[M.PL]	à.ART.DEF.PL	non_rom-PL	

« Nous on est des Roms/ qui/ on va/ on participe/ aux réunions/ non-roms. »

L'adjectif *gadžikane,* non-roms, est un adjectif relationnel dérivé de *gadžo, gadži*, Non-Rom. Il est question de réunions non-roms dans le syntagme (168), alors qu'il n'a pas été question de réunions roms auparavant. Dans ce type d'occurrence, le contraste qu'implique l'effet de sélection se fait *in absentia*, en l'absence (en discours ou dans la situation d'énonciation) des individus qui contrastent directement avec le référent. Une interprétation sélective fait le postulat suivant : le locuteur présuppose que je sais qu'il existe des réunions non-roms, en quoi elles consistent et en quoi elles diffèrent des réunions roms. Pourtant, cela n'a été nullement évoqué dans la conversation et ne le sera pas non plus par la suite.

Notons que ce syntagme présente non seulement une réduplication de l'article défini, mais aussi une réduplication de la préposition. Cela a de quoi étonner, car la préposition est, comme son nom l'indique, antéposée au syntagme nominal qu'elle introduit, donc au nom-tête – sinon il s'agit d'une postposition. Cela peut s'expliquer par le fait que la préposition *t*, à, est cliticisée à l'article défini *o* (à moins que ce ne soit l'inverse). Ils forment une seule syllabe et une seule unité : lorsque l'article est répété, la préposition est répétée avec lui. Ceci est symptomatique d'un haut degré de fusion, à l'instar des articles définis contractés avec la préposition *au* ou *du*, en français.

(169) *F24A, Bilisht, 27 juillet 2014*

O	*xabe*	*si*	***o***	***problemi***	***o***	***bar-o***.
ART.DEF.M.SG	nourriture(M.SG)	être.3.PRES	ART.DEF.M.SG	problème(M.SG)	ART.DEF.M.SG	grand-M.SG

« La nourriture c'est le grand problème. »[1]

L'analyse informationnelle et sémantique donne les résultats du Tableau 24.

Dans ces exemples, le référent de l'expression polydéfinie est accessible, tandis que les termes utilisés, noms et adjectifs, sont presque toujours nouveaux du point de vue lexical. Le syntagme constitue le plus souvent le focus de la phrase. L'expression peut s'interpréter soit grâce à un focus restrictif, souvent binaire, soit grâce à un focus emphatique non binaire. Le contraste se fait soit *in presentia*, soit *in absentia* – ce qui correspond en général à une interprétation plutôt à effet de sélection, ou à effet d'emphase.

[1] Cette locutrice vit dans un petit bidonville en périphérie de Bilisht, ville de montagne à huit kilomètres de la frontière grecque. Elle a des problèmes pour trouver à se nourrir (« On n'avait pas une miette de pain, pas de quoi manger »), pour se chauffer, pour s'habiller, pour se soigner. Les enfants ne sont pas inscrits à l'état civil, ne vont pas à l'école, etc.

Tableau 24 : Analyse informationnelle et sémantique des syntagmes DEF N DEF A à effet hybride

DEF N DEF A	Statut de l'information			Structure de l'information	Interprétation sémantique	Contraste
	référent du syntagme	**nom (lexicalement)**	**adjectif (lexicalement)**			
(164)	accessible-situationnel	nouveau	nouveau	focus de nouvelle information non sollicitée & focus-cadre de la phrase	intersection restriction pragmatique emphase	*in presentia*
(165)	accessible-général	nouveau	nouveau	topique-cadre	présupposé (N) sélection et focus contrastif *ou* focus assertif (A)	*in absentia*
(166)	accessible-général	donné	donné	focus de nouvelle information sollicitée	intersection restriction pragmatique opposition binaire emphase non binaire sur A superlatif absolu	*in presentia*
(167)	accessible-inférable (set relation) ou accessible-général	nouveau	nouveau	focus contrastif de remplacement	intersection restriction pragmatique emphase causale sur A	*in presentia*
(168)	accessible-situationnel	nouveau	nouveau	focus de nouvelle information non sollicitée	intersection restriction pragmatique opposition binaire emphase non binaire	*in absentia*
(169)	accessible-inférable	nouveau	nouveau	topique-propos	intersection restriction pragmatique superlatif relatif bilan emphase non binaire sur A	*in presentia*

2.1.2. DEF A DEF N

On rencontre quelques occurrences d'antéposition de l'adjectif dans le corpus. Cela n'a rien d'étonnant si l'on considère le grec moderne et les études qui lui sont consacrées, mais c'est inattendu pour les études sur le romani, qui ne mentionnent nulle part l'existence d'une possible structure DEF A DEF N. Le corpus des exemples (170) et (171), ainsi que le détail de l'analyse de chaque exemple, se trouve en section 1.2.5., p. 571.

(170) *M71A, Korçë, 28 juillet 2014*

Na	*mukh-l-e*	*la*	*o*	*čhaj-a*	***o***	***tikn-e***	***o***	***maksemi***.[1]
NEG	laisser-PRET-PL	3.SG.F.OBL	ART.DEF.PL	fille-F.PL	ART.DEF.PL	petit-PL	ART.DEF.PL	enfant(M.SG)

« Les filles, les petits enfants ne l'ont pas laissée [partir]. »

(171) *[Zamira][2], Korçë, 1er août 2014*

O	***barval-o***	***o***	***Skender***	*so*	*ker-d-as?*
ART.DEF.M.SG	riche-M.SG	ART.DEF.M.SG	Skender(M.SG)	INT	faire-PRET-3SG

« Skender le riche qu'est-ce qu'il a fait ? »

Dans l'exemple (171), on peut interpréter l'adjectif comme un cas d'*epithet*, adjectif de caractérisation ou étiquette (cf. section 2.4.4, p. 188) (Chatzikyriakidis 2014 : 1). Il s'agit d'une expression référentielle unique, où l'adjectif représente l'équivalent d'un titre apposé, lié au nom de la même manière que dans un nom composé. Alexiadou (2014 : 40-41) parlait, de son côté, d'interprétation *émotive* pour ce type d'adjectif : c'est exactement le cas ici, à ceci près que l'adjectif n'entraîne pas d'interprétation inverse de son sens propre. En grec comme en romani (et en italien lorsque l'adjectif est postposé, d'après Cinque 2010 : 87), ce type d'adjectif n'a pas d'interprétation restrictive et peut entrer dans un syntagme polydéfini.

L'analyse informationnelle et sémantique donne les résultats suivants (Tableau 25) :

Tableau 25 : Analyse informationnelle et sémantique des syntagmes DEF A DEF N

DEF A DEF N	Statut de l'information			Structure de l'information	Interprétation sémantique	Contraste
	référent du syntagme	nom (lexic.)	adjectif (lexic.)			
(170)	donné-actif	nouveau	nouveau	topique-propos	? intersection ? restriction pragmatique	*in absentia*

[1] Notons que le locuteur n'accorde pas le nom (qui est au singulier) et l'adjectif (qui est au pluriel). On attendrait *o tikne o maksemja* plutôt que *o tikne o maksemi*. Ce n'est pas la seule occurrence de morphologie inattendue chez ce locuteur : nous avons un autre exemple en (184), cf. note 1, p. 233.

[2] La locutrice a été anonymisée.

					expression ambiguë quasi nom composé ? propriété inhérente	
(171)	accessible-général	accessible-général	nouveau	topique-propos	caractérisation propriété inhérente	–

Si le référent de l'expression polydéfinie est nouveau ou accessible, les termes employés sont en revanche nouveaux du point de vue lexical, notamment l'adjectif. Le syntagme constitue le topique de la phrase. L'interprétation de l'expression se fait à l'intersection des ensembles de référents, grâce à une restriction pragmatique. La combinaison de l'adjectif et du nom peut former un quasi mot composé sur le plan sémantique.

2.1.3. Synthèse

2.1.3.1. Bilan

Le Tableau 116 constitue la synthèse des analyses présentées pour les exemples de syntagmes polydéfinis. On constate que dans nos exemples, le référent de DEF N DEF A ou DEF A DEF N est donné (sept occurrences) ou accessible (dix occurrences), rarement nouveau (une occurrence). Il n'est nouveau que dans l'exemple (162). Au contraire, l'adjectif est beaucoup plus souvent nouveau lexicalement (quatorze occurrences) que donné (quatre occurrences). C'est qu'il apporte une information nouvelle, une propriété qui permet l'identification du référent. Il est parfois donné, comme d'ailleurs le nom, qui peut être indifféremment donné, accessible ou nouveau.

L'analyse de la structure de l'information a montré que les syntagmes DEF N DEF A à effet de sélection et DEF A DEF N font le plus souvent partie du topique de la phrase (respectivement sept et deux occurrences) Les syntagmes DEF N DEF A à effet d'emphase font partie du focus de nouvelle information non sollicitée (les deux occurrences). Ces faits pourraient nous aider à interpréter les exemples « hybrides » de la section 2.1.1.3, p. 219. Les syntagmes DEF N DEF A y font partie du focus de nouvelle information non sollicitée (trois occurrences) ou de contraste (une occurrence). Dans les exemples (164), (166), (167) et (168), le syntagme joue le rôle de focus dans l'énoncé, ce qui fait pencher l'interprétation vers l'effet d'emphase, qui serait recherché par les locuteurs. Au contraire dans l'exemple (169), le syntagme joue le rôle de topique dans l'énoncé, ce qui fait pencher l'interprétation vers l'effet de sélection, qui serait recherché par la locutrice. Cela supposerait toutefois qu'il y ait un lien entre le type de focus étroit sur l'adjectif et la structure informationnelle de l'ensemble de l'énoncé. On constate qu'il n'y a pas d'incompatibilité entre le focus et le statut accessible ou donné de l'information transmise par le nom et/ou l'adjectif. Le focus ne correspond donc pas nécessairement à une information nouvelle. En réalité, seul l'exemple (162) présente un syntagme nominal qui soit à la fois sous focus et nouveau (référentiellement et lexicalement).

L'adjectif en DEF N DEF A fait toujours l'objet d'une interprétation intersective avec le nom : il sert à restreindre pragmatiquement, parmi une classe de référents, le choix à un référent en particulier (exemples (154), (155), (156), (159), (160) et (161)). Le processus à l'œuvre ici est le même qu'en grec moderne. L'adjectif transmet une propriété spécifique et contrastive (cf. notamment exemples (164), (166), (167), (168) et (169)), à l'exception de l'exemple (157) où c'est le nom qui apporte le sème (masculin ~ féminin) permettant l'identification. On a vu que la polydéfinitude pouvait servir d'artifice cognitif, de topic-shift ou, dans le cas de DEF A DEF N, de quasi mot composé. L'adjectif en DEF A DEF N peut soit faire l'objet d'une interprétation intersective, soit non.

Enfin, on a pu observer que l'interprétation à effet de sélection correspond à un contraste *in presentia* : les différents individus « concurrents », qui constituent l'ensemble de référents possible parmi lesquels va s'opérer la sélection, ont été mentionnés en discours ou bien sont présents dans la situation d'énonciation. Ce n'est pas le cas de l'interprétation à effet d'emphase, qui correspond à un contraste *in absentia*, qui s'effectue par exemple grâce à la connaissance générale du monde, au savoir que le locuteur présuppose commun et partagé avec l'interlocuteur.

2.1.3.2. Focus

L'affirmation la plus importante confirmée par l'analyse du corpus est la présence d'un focus sur l'adjectif : un focus étroit (*narrow focus*). Comme le grec moderne, selon Campos et Stavrou (2004 : 141-142), la polydéfinitude du romani est bâtie sur un contraste entre la partie nominale du syntagme, qui constitue le présupposé, et la partie adjectivale du syntagme, qui constitue le focus.

Les exemples (154) à (161) confirment là aussi l'affirmation de Campos et Stavrou (2004 : 147) pour qui l'article défini assume une fonction de focus contrastif et provoque un effet de sélection lors de l'interprétation du syntagme. Cependant, les exemples (162) et (163), certes moins nombreux, nous montrent que l'article défini peut assumer une autre fonction, qui rappelle le focus assertif, et provoquer un effet d'emphase. Les exemples (164) à (169) sont ambigus, le focus de la partie adjectivale pouvant être interprété de façon soit contrastive soit assertive, soit les deux à la fois.

Ainsi, les deux valeurs du focus sont les suivantes :

- La première vise à opérer un choix paradigmatique au sein d'un ensemble : ce choix présuppose l'existence d'entités alternatives au sein de l'ensemble et que l'énoncé ne serait pas valide si le choix était effectué en faveur d'une autre entité. L'adjectif ou le syntagme est mis en valeur dans une perspective de contraste, on baptisera donc *contrastive* cette première valeur. Elle relève de la fonction expressive du langage.

- La seconde vise à mettre en valeur l'adjectif ou le syntagme pour attirer l'attention, mais sans nécessairement le contraster à un autre référent possible. Cet usage ne présuppose pas qu'il existe d'autres membres alternatifs ou similaires au sein du même ensemble (l'exemple (162) ne suppose

pas qu'il y ait de mauvais films). On baptisera *assertive* cette deuxième valeur. Elle relève de la fonction conative du langage.

Le Tableau 26 résume ces résultats en présentant les différents types de focus des exemples, à la fois celui de la structure informationnelle de l'énoncé, et celui sur l'adjectif au sein du syntagme nominal.

Tableau 26 : Le focus en syntagme polydéfini

Structure	n°	Syntagme	Focus large (sur NP)		Focus étroit (sur A)		Pas de focus
			informatif	contrastif	informatif	contrastif	
DEF N DEF A	(154)	*o phral o vaver*		✓		✓	
	(155)	*mo phral o xurdo*				✓	
	(156)	*mi phen i/ i tikni*				✓	
	(157)	*mo phral o tikno*			✓		
	(158)	*o phral o baro*				✓	
	(159)	*i čhaj i vaver*				✓	
	(160)	*i škola i bari*				✓	
	(161)	*o phral o tikno*			✓	✓	
	(162)	*o filme o bare*	✓		✓		
	(163)	*i bira i purani*	✓		✓		
	(164)	*o lokali o baro*	✓		✓	✓	
	(165)	*o lokal o baro*			✓	(✓)	✓
	(166)	*amaro ćher o baro*	✓		✓	✓	
	(167)	*o čhave o bare*		✓	✓	✓	
	(168)	*to ćidipe to/ gadžikane*	✓		✓	✓	
	(169)	*o problemi o baro*			✓	✓	
DEF A DEF N	(261)	*o tikne o maksemi*					✓
	(171)	*o barvalo o Skender*			✓		

On constate que le type de focus présent sur l'adjectif et le type de focus dont fait l'objet l'ensemble du syntagme nominal ne sont pas nécessairement les mêmes. Il convient donc de dissocier les deux.

Comme l'interprétation du syntagme nominal est toujours intersective, on peut donner raison à Campos et Stavrou (2004 : 142) qui affirme : « *the definite article of the adjective in a polydefinite construction undertakes the same function as contrastive focus on the adjective in a monadic DP, the overall effect being one of contrast or selection among the referents of the denotation of the noun which, by not being stressed, constitutes the context or a contextually retrievable set* ». Le focus contrastif a pour fonction d'effectuer une opération de sélection. Cette opération a pour effet, en quelque sorte secondaire, d'éveiller l'attention de l'interlocuteur. Cet effet secondaire, devenu fonction secondaire, a pu, dans certains cas, motiver à lui seul l'usage de la focalisation. La fonction d'emphase serait alors comme une expansion de la fonction primaire de sélection. Mais nous avons vu que selon une autre hypothèse, la fonction primaire du focus est au contraire d'attirer l'attention (cf. section 3.3, p. 93). Ne pouvant trancher entre ces deux hypothèses, nous nous en tiendrons au

constat suivant : la partie adjectivale de la structure polydéfinie fait l'objet d'un focus qui a deux fonctions et qui peuvent aller de pair. L'existence de deux types de focus n'est donc pas à prendre comme une partition étanche des usages : il s'agit plutôt d'un continuum, avec des emplois concomitants.

Nous avons observé que la structure polydéfinie servait éventuellement à d'autres fins encore qu'à sélectionner et à attirer l'attention. Dans les exemples (158), (159), (160) et (261), elle permet au locuteur d'apporter une précision pour identifier le référent, et joue même presque le rôle d'une auto-correction en (158) et (261). Dans l'exemple (169), elle permet de récapituler le discours qui précède. Dans l'exemple (161), elle permet de structurer le récit en attirant l'attention de l'auditeur sur le changement de personnage principal et donc le changement de partie dans le récit. Ces fonctions annexes sont selon nous permises par la fonction assertive du focus.

2.1.3.3. Place de l'adjectif

Une question est soulevée par la découverte de DEF N DEF A en romani. L'ordre des mots importe-t-il : existe-t-il une différence sémantique et/ou fonctionnelle entre DEF A DEF N et DEF N DEF A ?

Il est difficile de se faire une idée précise à partir de deux exemples seulement : l'un semble un mot composé (exemple (261)), ce qui le rend très différent de DEF N DEF A… mais il joue le rôle d'apposition apportant une précision pour mieux identifier le référent, très exactement comme l'exemple (158). L'autre (exemple (171) ne semble pas présenter de différence sémantique ou fonctionnelle, puisque l'interprétation est intersective. Les études sur le grec fournissent très peu d'indices permettant de distinguer DEF A DEF N et DEF N DEF A. Campos et Stavrou (2004 : 147) en fournissent un concernant la prosodie :

> « *If the adjective precedes the noun, it is contrastively stressed and the adjective is interpreted under narrow focus, with emphatic stress being assigned to it by the Emphatic/Contrastive Stress Rule.* […] *Thus, we see that in a monadic construction, depending on context, either the adjective or the noun may carry contrastive stress, whereas in a polydefinite the adjective is always the one to be stressed, the noun providing only contextualized/old information*. » Campos et Stavrou (2004 : 147)

Cette assertion n'est pas valable pour notre corpus romani : dans les exemples (261) et (171), l'adjectif ne porte aucun accent emphatique. Le seul mot à porter un accent, l'accent nucléaire du syntagme, est le nom : *o tikne o makSEmi* et *o barvalo o SKENder*. Pour y voir plus clair, il nous faut étudier l'ordre des mots dans les syntagmes monodéfinis : existe-t-il une différence sémantique ou fonctionnelle entre DEF A N et DEF N A ?

2.2. Syntagmes monodéfinis

Le fil rouge qui guidera cette section est la distinction entre les deux types de position adjectivale établie par Cinque (2010, 2014), inspirée notamment de Sproat et Shih (1990) (cf. section 3.5, p. 100).

La position adjectivale reflète en effet une distinction entre modification directe et indirecte. Campos et Stavrou (2004 : 150) affirment que « *Greek displays both indirect (polydefinites) and direct (monadics) modification* ». Ce ne serait donc pas l'ordre des mots, mais la répétition ou non de l'article défini qui ferait le départ entre modification directe et indirecte.

2.2.1. DEF A N

Nous avons vu en section 1.4.3.1, p. 138 que la place canonique de l'adjectif est prénominale. La séquence A N représente l'ordre hérité des langues indiennes (Masica 1991 : 373) et celui le plus fréquent dans les différentes variétés de romani, notamment celles conservatrices (romani de Lituanie, Tenser 2005 : 54-55). Les exemples (172) à (176) sont extraits du corpus spontané. Le corpus dont ils sont issus, ainsi que le détail de l'analyse de chaque exemple, se trouve en section 1.2.6., p. 575.

(172) *F60A, Korçë, 21 juillet 2014*

Sa	***i***	***roman-i***	***čhib***	*jek*	*isi*
tout	ART.DEF.F.SG	romani-F.SG	langue(F.SG)	IND	être.3.PRES

« Toute la langue romani est une et une seule. »

(173) *M71A, Korçë, 28 juillet 2014*

Eh	*akowa*	*phira-d-e*	*otakal/*	*t-**o***	***puran-e***	***breš-a.***
eh	DEM	marcher-PRET-3PL	ainsi	à-ART.DEF.PL	ancien-PL	année-PL

« Eh comme ça, ils sont partis/ il y a longtemps. (litt. dans les anciennes années). »

(174) *M50A, Korçë, 5 août 2014*

I	***tik-ni***	***čhaj***	*ćer-d-om*	*la*	*othe/*	*Grećija.*
ART.DEF.F.SG	petit-F.SG	fille(F.SG)	faire-PRET-1SG	3.SG.F.OBL	là	Grèce

« La petite fille/ je l'ai faite là/ en Grèce. »[1]

(175) *M46A, Korçë, 22 juillet 2013*

Sa	*isi*	*jazmiši/*	*kat-**o***	***bar-o***	***Del.***
tout	être.3.PRES	écrit	venant_de-ART.DEF.M.SG	grand-M.SG	Dieu(M.SG)

« Tout est écrit/ par Dieu tout puissant. (litt. par Dieu grand) »

Le dieu auquel il est fait référence, et c'est pourquoi nous y avons mis une majuscule, est le dieu chrétien orthodoxe, qui a pour attribut d'être tout puissant. On pourrait objecter que Dieu fait partie de la connaissance du monde attendue, ce qui le rend accessible-général. Dans ce cas, l'adjectif *grand* transmet une propriété elle aussi accessible-général. L'adjectif *baro*, grand, transmet une propriété du référent qui lui est inhérente, sa puissance. Un dieu est forcément grand. En outre dans une

[1] Rien, dans la conversation, n'est dit du lieu de naissance des autres enfants. Le locuteur a quatre autres enfants, dont deux sont nés en Albanie d'une première femme, et deux sont également nés en Grèce, d'une deuxième femme.

perspective monothéiste, l'adjectif *grand* n'a pas une interprétation intersective avec le nom *Dieu* : il n'y a pas de restriction pragmatique puisqu'il n'y a qu'un seul dieu.

(176) *M71A, Korçë, 28 juillet 2014*

Pan-es-ć-e	*than-es-te*	*pij-as*	*sine/*	***o***	***šutal-o***	***thud.***
eau-M.SG.OBL-GEN-M.SG.OBL	lieu-M.SG.OBL-DAT	boire-1PL	être.3.PRET	ART.DEF.M.SG	acide-M.SG	lait(M.SG)

« Au lieu de l'eau on buvait/ du lait fermenté. »

La référence ne se calcule par la simple adjonction des sèmes de l'adjectif 'acide' et de ceux du nom 'lait' : il s'agit d'un type de boisson qui ne s'oppose pas à du lait non-acide, mais, en bloc, à de l'eau. Le référent est proche du kéfir ou du lait ribot (dont le sens représente plus que la simple somme de 'lait' et 'ribot'). La séquence A + N fonctionne comme un nom composé. En outre, le référent lait fermenté a un sens générique dans cet extrait.

(177) *F44AČ, Korçë, 1er août 2014*

O	***bilačh-o***	***manuš***	*ov-ela*	*vampiri.*
ART.DEF.M.SG	méchant-M.SG	homme(M.SG)	devenir-3SG.PRES	vampire

« Le méchant homme devient vampire. »

(178) *F65A, Korçë, 1er août 2014*

O	***lačh-e***	***manuša***	*dž-ala*	*leng-o*	*ogi*	*to*	*devel!*
ART.DEF.PL	bon-PL	homme(M.PL)	aller-3SG.PRES	POSS.3PL-M.SG	esprit(M.SG)	à.ART.DEF.M.SG	ciel(M.SG)

« Les gens (litt. hommes) gentils, leur âme va au ciel ! »

(179) *M46A, Korçë, 22 juillet 2013*

em	*ćer-ava*	*lafi*	***e***	***bar-e***	***doktor-en-sar.***
et	faire-1SG.PRES	parole	ART.DEF.PL.OBL	grand-PL.OBL	docteur-M.PL.OBL-INSTR

« […] et je discute avec les grands docteurs. »

L'analyse informationnelle et sémantique donne les résultats suivants (Tableau 27) :

Tableau 27 : Analyse informationnelle et sémantique des syntagmes DEF A N

DEF A N	Statut de l'information			Structure de l'information	Interprétation sémantique
	référent du syntagme	**nom (lexicalement)**	**adjectif (lexicalement)**		
(172)	donné	donné	nouveau ou accessible	topique-propos	intersection hypéronymie modification directe
(173)	nouveau	nouveau	nouveau	topique-cadre	sens général et non spécifique modification directe
(174)	accessible-inferable *set-relationship*	donné-inaccessible	nouveau	topique-propos	intersection restriction pragmatique focus étroit (A) modification indirecte
(175)	nouveau ou accessible-général	nouveau	nouveau	topique-cadre	interprétation hiérarchique propriété du référent qui lui est inhérente épithète de nature modification directe
(176)	donné	donné	donné	focus contrastif *subtype replacing*	propriété du référent qui lui est inhérente quasi nom composé modification directe
(177)	donné	donné	donné	topique-propos	intersection restriction pragmatique opposition binaire générique mariage sémantique contraste modification directe
(178)	nouveau	donné-inactif	nouveau ou accessible-inférable	topique-propos	
(179)	donné-inactif	donné-inactif	nouveau	focus de nouvelle information non sollicitée	propriété du référent qui lui est inhérente épithète de nature valeur emphatique, non contrastive modification directe

Dans ces exemples, le référent du syntagme est quasiment toujours donné ou accessible. Les termes sont indifféremment donnés ou nouveaux, en ce qui concerne le nom, et presque toujours nouveaux, en ce qui concerne l'adjectif. L'expression fait partie du topique de la phrase, dans la plupart des cas. L'adjectif est porteur d'une propriété inhérente au référent, qui s'interprète de manière hiérarchique. Il constitue quasiment un épithète de nature du nom, ou bien un nom composé avec lui. Ce sont des cas de modification directe du nom. Notons que lorsqu'il y a contraste, comme en (177) et (178), c'est de manière non spécifique. L'adjectif est mis au premier plan *avec* le nom, ce qui produit un effet de mariage sémantique (Nølke 2001 : 181). Il y a donc contraste, quand bien même il y a modification directe et il n'y a pas focus étroit sur l'adjectif.

2.2.2. DEF N A

La postposition de l'adjectif n'est pas l'emplacement canonique de l'adjectif, il s'agit d'une innovation dans l'histoire de la langue. Elle présente selon la littérature un caractère de « commentaire » (Boretzky 1993 : 41) ou sert une fonction pragmatique d' « *afterthought* » (Matras 2002 : 102). On peut considérer cette position comme marquée. Les exemples (180) à (186) sont extraits du corpus spontané. Le corpus dont ils sont issus, ainsi que le détail de l'analyse de chaque exemple, se trouve en section 1.2.7., p. 586.

(180) *F16AČ, Korçë, 22 juillet 2014*

Phirav-asa	***o***	***pet-e***	***bar-e***.
pétrir-1PL.PRES	ART.DEF.PL	boule_de_pâte-F.PL	grand-PL

« On pétrit les grosses boules de pâte. »

Contrairement à l'exemple (175) où l'adjectif était antéposé, l'adjectif *grand* n'est pas en (180) une propriété inhérente du référent puisque contrairement à Dieu, une boule de pâte n'a pas de taille unique. L'adjectif apporte ici un sème spécifique supplémentaire et sert à identifier le référent : la propriété qu'il transmet est présentée comme particulièrement importante, puisque la réussite de la procédure en dépend. Le syntagme en (180) est un exemple de modification indirecte car son interprétation est intersective et l'identification du référent se fait par une restriction pragmatique au sein des référents du nom.

(181) *M58MČ, Voskopojë, 15 juillet 2013*

ə/	***e/***	***čhav-e***	***bar-e***[1]	*si*	*le/*
euh	ART.DEF.M.SG.OBL	fils-M.SG.OBL	grand-M.SG.OBL	être.3.PRES	3PL

[1] Notons que ce syntagme est au cas oblique, tandis que dans la phrase suivante, bâtie sur le même modèle, le syntagme est au cas direct. Il y a donc variation morphologique pour le sujet de « avoir », littéralement « être à ». S'il est un pronom, il est toujours à l'oblique, s'il est un syntagme nominal, il est soit au direct soit à l'oblique.

ə	*duj*	*čhej-a*	*jek*	*čhavo.*
euh	deux	fille-F.PL	un	fils(M.SG)

« Euh/ le grand garçon il a/ euh deux filles un fils. »

(182) *M58MČ, Voskopojë, 15 juillet 2013*

ə	***o***	***čhavo***	***cikn-o/***	*murn-o*	*čhavo*	[...]
euh	3SG.M	fils(M.SG)	petit-M.SG	POSS1SG-M.SG	fils(M.SG)	
vo	*ə/*	*lje*	*si*	*le*	*jek*	*če.*
3SG.M	euh	3SG.M.OBL	être.3.PRES	3SG.M.OBL	un	fille(F.SG)

« Euh le petit garçon/ mon fils qui a/ euh/ trente cinq ans/ euh/ lui euh/ lui il a une fille. »

(183) *F16AČ, Korçë, 24 juillet 2014*

Sosće	*akale*	*ba/*	*bar-isajl-e*	*to/*	*sar*
parce_que	DEM.PL	gran/	grandir-PRET-3PL	à-ART.DEF.PL	comme
to	*Rom-a/*	*t-**o***	***amal-in-ja***	***vaver-a.***	
à-ART.DEF.PL	Rom-PL	à-ART.DEF.PL	ami-F-PL	autre-PL	

« Parce que ceux-ci ils ont/ ils ont grandi dans/ comme chez les Roms/ euh/ chez les autres amies. »

(184) *M71A, Korçë, 27 juillet 2013*

Bister-d-om	***odole***	***gil-ja***	***koraxan-e.***
oublier-PRET-1SG	DEM.F.SG	chanson-F.PL	musulman-PL

« Maintenant j'ai oublié ces chansons turques. »

Le locuteur a commencé par mentionner les chansons turques que chantaient les anciens[1], avant de mentionner le répertoire romani, et de conclure que les chansons turques étaient plus fréquemment chantées. Il a oublié les chansons turques anciennes, mais pas les autres.

(185) *M58MČ, Voskopojë, 14 juillet 2013*

*An-**o***	***foro***	***bar-o***	*Tirana/*	*but*	*tatipe.*
dans.ART.DEF.M.SG	ville(M.SG)	grand-M.SG	Tiranë	beaucoup	chaleur(M.SG)

« Dans la capitale à Tiranë/ il fait très chaud. »

Le locuteur réside justement dans la capitale. Mais cette conversation a lieu en plein été, le 14 juillet 2013, dans un café situé dans un bois dans les hauteurs de Moscopole/Voskopojë, ville de petite montagne. Un contraste existe, dans le contexte d'énonciation, entre *la grande ville* et le lieu où se tient la conversation. Impossible de savoir en revanche si le locuteur entend opposer la ville à la campagne (le café dans les bois) ou bien la métropole à la petite ville provinciale

[1] Notons un problème d'accord, dans le syntagme *ko šele brešengi gilja*, entre le nom (qui est au pluriel) et le modifieur, un groupe génitival (qui est au singulier). On attendrait en effet *ko šele brešenge gilja*. Nous rencontrons un problème de morphologie similaire dans l'exemple (261), cf. note 1, p. 223.

(Moscopole/Voskopojë)… donc de savoir si le focus à valeur contrastive porte sur l'ensemble N+A ou seulement sur A.

(186) *M68A, Voskopojë, 15 juillet 2013*

Kar-əna	*pes/*	*maškare*	*len-de/*	***pir-e***	***gad-a***	***roman-e.***
faire-3PL.PRES	3.REFL	entre	3.PL.OBL-DAT	POSS.3REFL-PL	vêtement-M.PL	rom-PL

« Ils se font eux-mêmes/ entre eux/ leurs habits roms. »[1]

Dans l'exemple (186), le groupe de Roms en question ne portent qu'un seul type de vêtement : l'adjectif « roms » n'opère donc pas de restriction au sein de l'ensemble {leurs vêtements}. Toutefois, le locuteur sait qu'il existe des vêtements « non-roms », que ce groupe de Roms pourrait choisir de porter occasionnellement. On peut penser qu'il les inclut inconsciemment dans l'ensemble plus général {vêtements}. l'adjectif « roms » n'opère pas de restriction au sein de l'ensemble {leurs vêtements}.

L'analyse informationnelle et sémantique donne les résultats du Tableau 28.

Dans ces exemples, le référent du syntagme est le plus souvent nouveau ou accessible par la connaissance du monde. Le nom est le plus souvent nouveau du point de vue lexical, tandis que l'adjectif est soit nouveau, soit donné. Le syntagme fait indifféremment partie du topique ou du focus de la phrase. Tous ces cas de postposition de l'adjectif s'interprètent comme de la modification indirecte. Le contraste s'effectue soit *in presentia*, soit *in absentia*.

[1] Cet exemple provient du même entretien que l'exemple (168). Le corpus transcrit et traduit se trouve p. 569.

Tableau 28 : Analyse informationnelle et sémantique des syntagmes DEF N A

DEF N A	Statut de l'information			Structure de l'information	Interprétation sémantique	Contraste
	référent du syntagme	**nom (lexicalement)**	**adjectif (lexicalement)**			
(180)	nouveau	nouveau	nouveau	focus de nouvelle information non sollicitée	intersection contraste focus étroit (A) à valeur contrastive implicite	*in absentia*
(181)	accessible-inférable (*set relation*)	donné-inactif	donné-inactif	topique-propos	intersection restriction pragmatique présupposé (N) focus étroit (A) à valeur contrastive modification indirecte	*in presentia*
(182)	accessible-inférable (*set relation*)	donné	donné-inactif	topique-propos	intersection restriction pragmatique présupposé (N) focus étroit (A) à valeur contrastive modification indirecte	*in presentia*
(183)	nouveau	nouveau	nouveau	focus de nouvelle information non sollicitée	intersection restriction pragmatique présupposé (N) focus étroit (A) à valeur contrastive modification indirecte	*in absentia*
(184)	donné (-inactif)	donné (-inactif)	donné (-inactif)	focus de nouvelle information sollicitée	intersection restriction pragmatique présupposé (N) focus étroit (A) à valeur contrastive adjectif spécifique modification indirecte	*in absentia / presentia*
(185)	nouveau ou accessible-général	nouveau	nouveau	topique-cadre	intersection restriction pragmatique présupposé (N) focus étroit (A) à valeur contrastive adjectif spécifique modification indirecte	*in presentia*
(186)	nouveau	nouveau	nouveau ou accessible-inférable	focus de nouvelle information non sollicitée	intersection ? restriction pragmatique présupposé (N) focus étroit (A) à valeur contrastive modification indirecte	*in absentia*

2.2.3. Synthèse

2.2.3.1. Bilan

Le Tableau 117 présente la synthèse des analyses. On constate que dans nos exemples, le référent de DEF A N peut-être donné (six occurrences), accessible (deux occurrences), ou nouveau (cinq occurrences). Le référent de DEF N A peut-être donné (une occurrence), accessible (deux occurrences), ou nouveau (cinq occurrences). Il est un peu plus souvent donné/accessible (huit occurrences) que nouveau (cinq occurrences) en DEF A N – c'est le contraire en DEF N A (respectivement trois et cinq occurrences). Le statut de l'information ne semble donc pas très pertinent pour les exemples présentés. L'adjectif, quant à lui, est beaucoup plus souvent nouveau lexicalement (quatorze occurrences) que donné (six occurrences) ou accessible (une occurrence) : il apporte une information nouvelle, une propriété qui permet l'identification du référent.

Les syntagmes DEF A N font le plus souvent partie du topique de la phrase (onze occurrences) que du focus de la phrase (deux occurrences), au contraire des syntagmes DEF N A, qui font le plus souvent partie du focus (cinq occurrences) que du topique (trois occurrences). Il y a là un point commun entre DEF A N et DEF A DEF N d'une part, et un point commun entre DEF N A et DEF N DEF A d'autre part.

L'adjectif antéposé transmet souvent une propriété non-spécifique (exemples (172) et (173)), générique (exemple (176) voire inhérente (exemples (175), (176) et (179)), ou une restriction (exemples (172) et (174)). Il s'agit de modification directe. Il peut parfois transmettre une propriété spécifique (exemple (174). Au contraire, l'adjectif postposé transmet une propriété spécifique (tous les exemples) et une restriction (tous les exemples), éventuellement contrastive (tous les exemples sauf (181) et (182)). L'adjectif reçoit un focus étroit. Il s'agit de modification indirecte.

On a pu constater que le référent des syntagmes DEF N A fait l'objet d'un contraste, soit *in presentia* (en majorité), soit *in absentia*.

2.2.3.2. Focus

Un adjectif peut faire partie d'un syntagme qui constitue le focus de la phrase. Nous avons vu aussi que la partie adjectivale (accompagnée ou non d'un article) du syntagme jouait le rôle de focus si elle est postposée, tandis que la partie nominale jouait le rôle de présupposé. Le Tableau 29 présente les types de focus dans les exemples monodéfinis.

Tableau 29 : Focus en syntagme monodéfini

Structure	n°	Syntagme	Focus large (sur NP)		Focus étroit (sur A)		Pas de focus
			informatif	contrastif	informatif	contrastif	
DEF A N	(172)	*sa i romani čhib*					✓

Structure	n°	Syntagme	Focus large (sur NP)		Focus étroit (sur A)		Pas de focus
			informatif	contrastif	informatif	contrastif	
	(173)	*to purane breša*					✓
	(174)	*i tikni čhaj*					✓
	(175)	*kato baro Del*					✓
	(176)	*o šutalo thud*		✓			
	(177)	*o bilačho manuš*				✓	✓
	(178)	*o lače manuša*				✓	✓
	(179)	*e bare doktorensar*	✓				
DEF N A	**(180)**	*o pete bare*	✓			✓	
	(181)	*e čhave bare*				✓	
	(182)	*o čhavo cikno*				✓	
	(183)	*to amalinja vavera*	✓		✓		
	(184)	*odole gilja koraxane*	✓		✓		
	(185)	*ano foro baro*				✓	
	(186)	*pire gada romane*	✓			✓	

On constate que, comme dans les syntagmes polydéfinis, le type de focus présent sur l'adjectif (en DEF N A) et le type de focus dont fait l'objet l'ensemble du syntagme nominal ne sont pas nécessairement les mêmes. Il convient donc de dissocier les deux. C'est pourquoi un syntagme DEF A N, tel que celui de l'exemple (176), qui ne présente pas de focus étroit sur l'adjectif, peut faire partie du focus large de la phrase.

Ceci n'implique pas pour autant que l'adjectif soit sous focus étroit chaque fois qu'il contraste sémantiquement avec un autre. Le contraste peut avoir lieu alors que l'adjectif est antéposé et qu'il y a donc modification directe. Ces cas s'expliquent grâce à Nølke (2001 : 180-181) cité plus haut (section 2.2.3.2, p. 235) : l'adjectif est focalisé en bloc avec son substantif.

2.2.3.3. Place de l'adjectif

La position de l'adjectif permet de présenter l'information de manière différente :

- La place antéposée permet un emploi peu ou pas spécifique de l'adjectif, voire générique. Elle permet de délivrer une propriété inhérente au nom ou de référer de manière hyperonymique. L'adjectif fait partie intégrante de l'expression référentielle globale, jusqu'à pouvoir former un mot composé avec le nom.

- La place postposée permet un emploi spécifique de l'adjectif. La référence de l'expression globale se trouve à l'intersection des référents du nom et de l'adjectif et l'adjectif permet de le sélectionner. L'adjectif reçoit alors soit un focus contrastif, soit un focus assertif.

On voit que les deux rôles sont conçus en opposition et que la partition se fait de façon binaire via des positions prévues par la syntaxe. La première manière de présenter l'information rappelle le rôle de présupposé, la seconde rappelle le rôle de focus (Nølke 2001 : 91). C'est le cas pour tous les adjectifs en position postposée. La présence de focus est en revanche indépendante de la fonction de focus ou de topique que joue par ailleurs le syntagme nominal au sein de l'énoncé, ce qui infirme l'hypothèse faite précédemment.

L'analyse de Campos et Stavrou (2004) ne vaut donc pas seulement pour les polydéfinis grecs : elle vaut pour les adjectifs postposés en romani, monodéfinis (DEF N A) ou polydéfinis (DEF N DEF A). Il est plus difficile de s'en apercevoir en grec, où une structure telle que DEF N A n'existe pas, à moins de former une prédication nominale. En effet, dans les exemples en DEF N A comme dans les exemples en DEF N DEF A, l'adjectif a une fonction particulière : celle d'identifier le référent. La propriété permettant de l'identifier est livrée par l'adjectif, parce qu'il apporte des *précision*s, un plus haut degré de spécificité qui permettent cette identification.

Voici une phrase présentant l'adjectif dans les deux positions :

(187) *M46A, Korçë, 8 août 2014*

Te	*nik-av*	*jek*	*ə/*	*jek*	*ə/*
COMPL	sortir-1SG.PRES	IND	euh	IND	euh
ə	***jek***	***ruba/***	***karakteristik-o***	***romandikan-o/***	***t-o***
euh	IND	?(F.SG)	caractéristique-M.SG	romani-M.SG	à-ART.DEF.M.SG
vaćeribe	***řoman-o***	*na*	*čhiban-i*	*čhib.*	
parler(M.SG)	romani-M.SG	NEG	albanais-F.SG	langue(F.SG)	

« Pour que je produise un euh/ un euh/ euh un (?)[1]/ caractéristique romani/ dans le parler romani pas en langue albanaise. »

Le référent « le parler romani » est nouveau dans la conversation. L'adjectif *řomano* est cependant accessible-inférable lexicalement, puisque son synonyme *romandikano* est prononcé juste avant lui. Le référent « la langue albanaise » est nouveau dans la conversation. Le nom *čhib*, langue, est accessible-inférable puisqu'il a déjà été question de parler, mais il est nouveau lexicalement. L'adjectif *čhibano*, albanais est nouveau lexicalement. Les syntagmes sont tous deux en focus contrastif du type *subtype selection* chez (Götze et al. 2007).

La référence se calcule par l'adjonction des sèmes de l'adjectif et de ceux du nom : les référents des expressions globales sont à l'intersection de l'ensemble {parler} ou {langue} et de l'ensemble

[1] Je suppose qu'en tant que journaliste, il parle ici d'un film ou d'un documentaire.

{entités romani} ou {entités albanaises}. Les deux référents « le parler romani » et « la langue albanaise » font l'objet d'une opposition binaire. Un focus étroit est porté sur l'adjectif dans les deux cas, cependant le locuteur choisit de mettre en avant *romani* plutôt qu'*albanaise*, qui fait l'objet d'une négation. Le locuteur postpose l'adjectif afin de mettre en avant la propriété de l'adjectif *romano* : le fait que le documentaire soit en langue romani pourrait étonner certains, d'où sa mise en valeur syntaxique et sémantique. Au contraire, l'antéposition de l'adjectif semble refléter cette mise au second plan de l'albanais, comme si la syntaxe venait ici appuyer le sens de la phrase.

Une autre hypothèse serait que le locuteur veuille souligner le parallèle avec le syntagme précédent, *jek ruba/ karakteristiko romandikano*, un (?)/ caractéristique romani, où les adjectifs *karakteristiko* et *romandikano* sont postposés – comme le plus souvent dans les syntagmes indéfinis présentant une information nouvelle (cf. Boretzky 1993 : 41). Cela expliquerait pourquoi il conserve cette postposition de l'adjectif dans le syntagme défini qui suit, *to vaćeribe řomano*, dans le parler romani. Le parallèle structurel entre *jek ruba karakteristiko romandikano* et *vaćeribe řomano* souligne sémantiquement la cohérence du projet : produire un film sur le thème des Roms dans la langue des Roms.

Notons que dans la phrase suivante (cf. corpus p. 595), le locuteur prononce le syntagme *sa sa řomani čhib*, tout tout en langue romani, qui ne présente pas d'article défini mais où l'adjectif est antéposé (exactement comme *na čhibani čhib*, pas en langue albanaise). Le référent « la langue romani » est donné-actif, le nom *čhib*, langue, et l'adjectif *řomani* sont donné-actif lexicalement. Le syntagme *sa sa řomani čhib*, tout tout en langue romani, constitue le focus de la phrase… ce qui peut sembler contradictoire avec l'antéposition de l'adjectif. Celle-ci peut être dû au statut donné du référent et de l'information portée par le syntagme, ainsi qu'au caractère redondant de l'information, déjà donnée dans la phrase précédente. Une autre hypothèse serait, enfin, la variation dialectale.

C'est également la focalisation qui explique pourquoi des adjectifs ordinaux tels que *deuxième*, *dernier*, sont toujours antéposés. « On a souvent noté que l'adjectif s'approche des déterminants et des quantificateurs dans l'antéposition (la dernière séance, pure bêtise, la première scène). Or ni la détermination ni la quantification ne servent à sélectionner des référents à l'intérieur d'un paradigme, mais par contre à ajouter des propriétés au référent déjà sélectionné. » (Nølke 2001 : 186). C'est également le cas en romani :

(188) *M68A, Voskopojë, 15 juillet 2013*

ə/	***o***	***jekto***	***kongreso***	*ondil-o/*	*Londra*
euh	ART.DEF.M.SG	premier	congrès(M.SG)	être.PRET-3.SG	Londres

« Euh/ le premier congrès a eu lieu/ à Londres […]. »

L'adjectif *jekto*, premier, est un numéral ordinal issu de *jek*, un et de *-to*, ième. Il est ici à sa place habituelle, antéposé au nom. Il est référentiellement et lexicalement donné-actif, puisqu'il vient d'être mentionné dans la phrase précédente – la pause visant à aider le locuteur à se remémorer la date exacte du congrès en question. Le syntagme constitue le topique-propos de l'expression *ondilo/*

Londra, a eu lieu à Londres. On constate qu'au contraire la plupart des autres numéraux ordinaux sont postposés au nom dans cet extrait :

- *kongreso o jekto e Romengo*, congrès premier des Roms
- *ko kongresi/ ko pandž/*, au cinquième/ congrès, qui ressemble à un syntagme polydéfini, où l'adjectif est remplacé par un numéral
- *kongreso i/ šov/*, le congrès ès/ six/, où le locuteur hésite entre les formes phonétiques *kongreso* (plus typique de la *khetani čhib*, langue de rassemblement de l'élite rom européenne, cf. section 1.3.2, p. 205) et *kongresi* (plus typique du romani d'Albanie) pour le mot « congrès », qui est un emprunt.
- *kongreso efta/ e Romengo*, le congrès sept/ des Roms
- *kongresi jekto kongresi e Romengo*, le congrès premier congrès des Roms est ambigu, car l'adjectif *jekto*, premier, est à la fois postposé à *kongresi* qui précède et à *kongresi* qui suit.

On peut noter que soit ces syntagmes font partie du focus (*large*) de l'énoncé, soit le numéral postposé reçoit un focus *étroit* au sein du syntagme nominal, puisque le locuteur est en train d'énumérer les congrès en égrenant leurs numéros. En revanche, l'exemple (188), le premier congrès, constitue le début de la liste, le locuteur n'est donc pas encore en train de contraster les congrès les uns avec les autres. S'il y a contraste, il est implicite et par anticipation, pour le locuteur, avec ce qui va suivre.

2.3. Synthèse générale

On remarque que tous les syntagmes de ce corpus ne présentent qu'un seul adjectif.

2.3.1. Type d'adjectif

• **Poids** : On a vu en section 2.2.2, p. 74, que le poids respectif du nom et de l'adjectif jouerait un rôle : l'élément le plus lourd tend à la postposition. On constate l'absence d'adjectifs quadrisyllabiques antéposés dans notre corpus – mais cela peut être dû à la moindre fréquence dans la langue des adjectifs quadrisyllabiques, et à la moindre quantité d'adjectifs antéposés dans notre corpus (Tableau 30). En dehors de l'absence de noms monosyllabiques précédant leur adjectif en syntagme monodéfini, le poids du nom ne semble pas jouer de rôle quant à sa position dans le syntagme (Tableau 31).

Tableau 30 : Poids et place de l'adjectif : quelques exemples du corpus spontané

Adjectifs	DEF A N	DEF N A	DEF A DEF N	DEF N DEF A
2 syllabes	*o baro phral* le grand frère	*e čhave bare* les grands enfants	*o tikne o maksemi* les petits enfants	*o čhave o bare* les grands enfants
3 syllabes	*to purane breša*	*pire gada romane*	*o barvalo o Skender*	*i bira i purani*

	dans les anciennes années	leurs habits roms	Skender le riche	l'ancienne brasserie
4 syllabes		(*odole gilja koraxane*) ces chansons turques		*to ćidipe to/ gadžikane* au rassemblement/ non-rom

Tableau 31 : Poids et place du nom : quelques exemples du corpus spontané

Noms	DEF A N	DEF N A	DEF A DEF N	DEF N DEF A
2 syllabes	*o vaver phral* l'autre frère		*o barvalo o Skender* Skender le riche	*o phral o vaver* l'autre frère
3 syllabes	*o bilačho manuš* le méchant homme	*o čhavo cikno* le petit enfant	*o tikne o maksemi* les petits enfants	*i škola i bari* la grande école
4 syllabes	*o jekto kongreso* le premier congrès	*to vaćeribe řomano* dans le parler romani		*to ćidipe to/ gadžikane* au rassemblement/ non-rom

• **Origine** : Du point de vue étymologique, les adjectifs du corpus sont tous hérités du fonds indien. Les noms sont aussi bien hérités (*phen*, sœur) qu'empruntés (*problemi*, problème), aussi bien communs que propres (*Skender*).

• **Syntaxe** : Dans le corpus polydéfini, tous les adjectifs sont prédicatifs, à l'exception de *vaver*, autre.

• **Sémantique** : Tous les adjectifs des syntagmes polydéfinis sont restrictifs, à l'exception de l'exemple (157) « petit », (163) « ancienne », (171) « riche » et éventuellement de (261) « petits ». Deux adjectifs peuvent s'interpréter comme lexicalisés avec le nom : (160) « grande » et (261) « petits ». Un adjectif est relationnel, (168) « non-roms ».

2.3.2. Modification directe et indirecte

Les exemples du corpus monodéfini montrent que la « géographie » du syntagme nominal n'est pas seulement une question d'emplacement canonique ou non. Il s'agit de différents types de modification qui se répartissent comme indiqué dans le Tableau 32 : la position antéposée est celle de la modification directe, tandis que la position postposée est celle de la modification directe ou, plutôt, indirecte.

Qu'en est-il du corpus polydéfini ? On voit, avec le Tableau 33, que le romani semble présenter les mêmes caractéristiques que le grec moderne selon : la polydéfinitude est une structure de modification indirecte. Nous posons cependant l'hypothèse que ce n'est pas nécessairement la polydéfinitude en soi, la présence multiple de l'article défini, qui rend la modification directe, mais la position de l'adjectif. Selon cette hypothèse, DEF N DEF A relèverait de la modification indirecte, mais DEF A DEF N relèverait de la modification directe, puisque l'adjectif y est antéposé. Pour cela, il nous faut expliquer les exemples (171) et (174), modifications indirectes que l'on attendrait directes. Revenons à la question du focus.

Tableau 32 : Modification directe et indirecte en syntagme monodéfini

Structure	n°	Syntagme	Modification directe	Modification indirecte
DEF A N	**(172)**	*sa i romani čhib*	✓	
	(173)	*to purane breša*	✓	
	(174)	*i tikni čhaj*		✓
	(175)	*kato baro Del*	✓	
	(176)	*o šutalo thud*	✓	
	(177)	*o bilačho manuš*	✓	
	(178)	*o lače manuša*	✓	
	(179)	*e bare doktorensar*	✓	
	(188)	*o jekto kongreso*	✓	
DEF N A	**(180)**	*o pete bare*		✓
	(181)	*e čhave bare*		✓
	(182)	*o čhavo cikno*		✓
	(183)	*to amalinja vavera*		✓
	(184)	*odole gilja koraxane*		✓
	(185)	*ano foro baro*	(✓ ?)	✓
	(186)	*pire gada romane*		✓
	(187)	*to vaćeribe řomano*		✓

Tableau 33 : Modification directe et indirecte en syntagme polydéfini

Structure	n°	Syntagme	Modification directe	Modification indirecte
DEF N DEF A	**(154)**	*o phral o vaver*		✓
	(155)	*mo phral o xurdo*		✓
	(156)	*mi phen i/ i tikni*		✓
	(157)	*mo phral o tikno*		✓
	(158)	*o phral o baro*		✓
	(159)	*i čhaj i vaver*		✓
	(160)	*i škola i bari*		✓
	(161)	*o phral o tikno*		✓
	(162)	*o filme o bare*		✓
	(163)	*i bira i purani*		✓
	(164)	*o lokali o baro*		✓
	(165)	*o lokal o baro*		✓
	(166)	*amaro ćher o baro*		✓
	(167)	*o čhave o bare*		✓
	(168)	*to ćidipe to/ gadžikane*		✓
	(169)	*o problemi o baro*		✓
DEF A DEF N	**(261)**	*o tikne o maksemi*	✓	
	(171)	*o barvalo o Skender*		✓

2.3.3. Place de l'adjectif

On a vu qu'il existe deux emplacements dans le syntagme nominal du romani d'Albanie. La zone antéposée au nom permet de mettre le modifieur en position thématique. C'est la zone de la modification directe. La zone postposée au nom permet de mettre le modifieur en position rhématique. C'est la zone de la modification indirecte. Le romani partage cette caractéristique avec le grec, l'aroumain et l'italien, selon Campos et Stavrou (2004). Cela supposent trois présupposés :

- Le syntagme a une structuration interne : la place du modifieur fait sens, et un changement de l'ordre des constituants entraîne un changement sémantique.
- Les deux positions de l'adjectif (antéposé et postposé) ont les mêmes valeurs en syntagme polydéfini et en syntagme monodéfini.
- Cette structuration interne n'a pas nécessairement de lien avec la structure informationnelle de l'énoncé : un adjectif peut être sous focus au sein d'un syntagme nominal, qui ne joue pas, lui, le rôle de focus dans l'énoncé.

La différence entre DEF A DEF N et DEF N DEF A est la même qu'entre DEF A N et DEF N A : l'adjectif n'est pas mis sous focus dans le premier cas, ou bien il l'est ensemble avec le nom, tandis que l'adjectif est mis sous focus dans le deuxième cas. Le contenu sémantique de l'adjectif est laissé dans l'ombre (les sèmes activés sont moins nombreux) en DEF A (DEF) N qu'en DEF N (DEF) A où il est mis en lumière. Cela confirme Campos et Stavrou (2004 : 156), selon lesquels la structure DEF N DEF A assigne une propriété temporaire au référent du nom, tandis que la structure DEF A N transmet une propriété permanente, inhérente.

Cela n'est toutefois pas, selon nous, dû à la polydéfinitude, donc à l'ajout d'un article défini, mais à la position de l'adjectif. La place antéposée est propre à transmettre une propriété permanente, tandis que celle postposée est propre à transmettre une propriété temporaire. Cette hypothèse explique l'apparente contradiction, chez Campos et Stavrou (2004 : 145-146), entre le fait que les syntagmes polydéfinis transmettent à la fois une propriété permanente (leur exemple 18 : *To kalo to palikari kseri ki alo monopati, The good guy can find his way*) et temporaire (leur exemple 20a : *I asterismi i orati perilamvanun ti Megali Arkto, The visible constellations include the Big Ursus*). C'est que leur exemple 18 de propriété permanente est une structure DEF A DEF N, tandis que leur exemple 20a de propriété temporaire est une structure DEF N DEF A.

On comprend mieux pourquoi DEF A DEF N est très rare en romani. L'adjectif est mis en position non focale, de présupposé, sans focus étroit… mais est accompagné d'un article défini qui le focalise. Une telle construction est presque oxymorique. Il s'agit donc toujours de cas très particulier. En (261), DEF A DEF N permet de combiner l'impératif d'identification avec un adjectif placé en position thématique, lié au nom de façon étroite à la manière d'un nom composé. En (171), le syntagme nominal *o barvalo o Skender*, Skender le riche, constitue une critique adressée au beau-frère et un

reproche pour sa pingrerie. Si elle avait postposé l'adjectif, elle aurait attiré l'attention sur la propriété de la richesse. La focalisation étant le résultat d'un choix paradigmatique, elle aurait rendu la charge polémique encore plus forte (« parmi les propriétés de mon beau-frère, je sélectionne sa richesse. »). Antéposer l'adjectif rend la critique plus mesurée ou diplomate. La locutrice fait semblant de n'avoir effectué aucun choix, comme si la richesse était inhérente au beau-frère. On peut aller jusqu'à dire que la richesse du beau-frère, posée comme une propriété permanente de celui-ci, rend son comportement encore plus mesquin aux yeux de la locutrice. L'attaque, si discrète que la rende la syntaxe, en devient peut-être plus âpre.

Les exemples montrent que la place de l'adjectif est en lien direct avec son sens : la position postposée (DEF N A et DEF N DEF A) permet de souligner les sèmes spécifiques de l'adjectif, tandis que la position antéposée (DEF A N et DEF A DEF N) les atténue (Nølke 2001 : 180). La fonction de la postposition est donc d'apporter une précision sémantique permettant de mieux identifier le référent du syntagme nominal. Plus un adjectif comporte de sèmes spécifiques, moins son extension est grande, moins il comporte de sèmes spécifiques, plus son extension est grande (Nølke 2001 : 181-182). Un adjectif de petite extension, lorsqu'il est antéposé, perd ses sèmes spécifiques et gagne en extension. Cela permet de lui attribuer diverses valeurs, des connotations et des effets de style. La position de focus requiert l'existence de sèmes spécifiques qui distinguent des éléments similaires et réduisent par là le nombre de référents possibles à identifier. Au contraire, la position de thème requiert moins de sèmes : elle n'a pas pour rôle principal l'identification du référent. L'emplacement de l'adjectif dépend de sa sémantique : plus un adjectif est spécifique, plus il a de chance d'être postposé. Un adjectif de sens général sera plus facilement antéposé, ce qui est la place la moins spécifique.

Enfin, ce corpus ne présente pas de succession d'adjectifs épithètes.

2.3.4. Place de l'adjectif et variation dialectale

Le Tableau 34 présente le profil des personnes ayant prononcé les exemples polydéfinis.

Tableau 34 : Profil des locuteurs des syntagmes polydéfinis

Structure	n°	Syntagme	Locuteur	Genre	Age[1]	Variété
DEF N DEF A	**(154)**	*o phral o vaver*	F60A	F	60	arli
	(155)	*mo phral o xurdo*	M18Č	M	18	čergar I
	(156)	*mi phen i/ i tikni*	M24M	M	24	arli
	(157)	*mo phral o tikno*	M24M	M	24	arli
	(158)	*o phral o baro*	M71A	M	71	arli
	(159)	*i čhaj i vaver*	F38A	F	38	arli

[1] A l'époque de l'enregistrement (2013 ou 2014).

Structure	n°	Syntagme	Locuteur	Genre	Age[1]	Variété
	(160)	*i škola i bari*	M23M	M	23	mečkar
	(161)	*o phral o tikno*	M46A	M	46	arli
	(162)	*o filme o bare*	F16AČ	F	16	arli & čergar I
	(163)	*i bira i purani*	M46A	M	46	arli
	(164)	*o lokali o baro*	F16AČ	F	16	arli & čergar I
	(165)	*o lokal o baro*	F16AČ	F	16	arli & čergar I
	(166)	*amaro ćher o baro*	F16AČ	F	16	arli & čergar I
	(167)	*o čhave o bare*	F66M	F	66	mečkar
	(168)	*to ćidipe to/ gadžikane*	M68A	M	2013	arli
	(169)	*o problemi o baro*	F24A	F	24	arli
DEF A DEF N	**(261)**	*o tikne o maksemi*	M71A	M	71	arli
	(171)	*o barvalo o Skender*	[Zamira]	F	xx	arli & čergar I

Ces résultats n'ont bien sûr pas force de loi, puisqu'ils ne sont représentatifs que des personnes que j'ai pu côtoyer dans ma vie en immersion dans la communauté arli de Korçë et lors des voyages effectués avec mon hôte. On constate pour ce corpus que :

- Les femmes comme les hommes pratiquent la polydéfinitude. Elle n'est donc pas spécifique à un groupe ou à un autre.
- Des personnes de tous âges pratiquent la polydéfinitude (de 16 à 71 ans). Elle n'est donc pas spécifique à une classe d'âge ou à une autre… à l'exception des moins de 16 ans, avec lesquels j'ai moins conversé et pour lesquels je n'ai pas de données ici.
- Des locuteurs de toutes les variétés que j'ai pu rencontrer pratiquent la polydéfinitude : arli (la majorité ici, puisque c'est cette communauté qui m'a accueillie), čergar I et mečkar (une minorité de personnes ici), ainsi que des personnes parlant plusieurs variétés (cf. section 1.3.2, p. 205).

Il semble que la polydéfinitude soit pratiquée par tous les locuteurs du romani d'Albanie, à l'exception des plus jeunes. Je n'ai cependant pas pu collecter de données auprès des locuteurs de deux variétés très minoritaires, bamlija et kurtofi. Le Tableau 35 présente le profil des personnes ayant prononcé les exemples monodéfinis.

Tableau 35 : Profil des locuteurs des syntagmes monodéfinis

Structure	n°	Syntagme	Locuteur	Genre	Age[1]	Variété
DEF A N	**(172)**	*sa i romani čhib*	F60A	F	60	arli
	(173)	*to purane breša*	M71A	M	71	arli
	(174)	*i tikni čhaj*	M50A	M	50	arli
	(175)	*kato baro Del*	M46A	M	46	arli

[1] A l'époque de l'enregistrement (2013 ou 2014).

Structure	n°	Syntagme	Locuteur	Genre	Age[1]	Variété
	(176)	*o šutalo thud*	M71A	M	71	arli
	(177)	*o bilačho manuš*	F44AČ	F	44	čergar I & arli
	(178)	*o lače manuša*	F65A	F	65	arli
	(179)	*e bare doktorensar*	M46A	M	46	arli
	(188)	*o jekto kongreso*	M68A	M	68	arli
DEF N A	**(180)**	*o pete bare*	F16AČ	F	16	arli & čergar I
	(181)	*e čhave bare*	M58MČ	M	58	mečkar & čergar I
	(182)	*o čhavo cikno*	M58MČ	M	58	mečkar & čergar I
	(183)	*to amalinja vavera*	F16AČ	F	16	arli & čergar I
	(184)	*odole gilja koraxane*	M71A	M	71	arli
	(185)	*ano foro baro*	M58MČ	M	58	mečkar & čergar I
	(186)	*pire gada romane*	M68A	M	68	arli
	(187)	*to vaćeribe řomano*	M46A	M	46	arli

On constate pour ce corpus que :

- Les syntagmes DEF A N sont presque tous prononcés par des locuteurs arli. Ces personnes ont toutes plus de 46 ans.
- Les syntagmes DEF N A sont prononcés par des locuteurs de différentes variétés et de tous âges (16 à 71 ans). Le corpus ne présente aucun locuteur de moins de 16 ans.

Il semble que l'antéposition de l'adjectif soit pratiquée exclusivement par les locuteurs arli et d'un certain âge. Cela signifierait que le *mapping* modification directe / indirecte sur l'emplacement antéposé / postposé au nom est beaucoup plus systématique dans les variétés mečkar et čergar I qu'en arli. Les locuteurs arli antéposent l'adjectif beaucoup plus fréquemment que les locuteurs d'autres groupes, même dans les cas de modification indirecte ou de focus, où l'on attendrait donc un adjectif postposé. On comprend mieux, alors, que l'exemple (174) soit une modification indirecte malgré l'antéposition de l'adjectif : il est prononcé par un arliophone. Au contraire, les locuteurs mečkar et čergar I postposent quasiment toujours l'adjectif, à tel point qu'il est plus aisé de trouver un exemple de DEF N A chez un Arli qu'un exemple de DEF N A chez un Mečkar ou un Čergar I.

Voici quelques autres occurrences d'antéposition « inattendue » de l'adjectif effectuée par des locuteurs arli. Dans les exemples (189) et (190), le locuteur est un homme de 27 ans – le seul de moins de 46 ans dont nous disposions pour le corpus de DEF A N.

(189) *M27A, Korçë, 5 août 2014*

Si/	*po*	*bar-e*	*man-dar*	***o***	***bar-o***
être.3.PRES	COMP	grand-PL	SG.OBL-ABL	ART.DEF.M.SG	grand-M.SG
phral	*isi*	*sarandupandž*	*e*	*breš-en-g-o*	

frère(M.SG) être.3.PRES quarante-cinq ART.DEF.PL.OBL année-M.PL.OBL-GEN-M.SG

« Ils sont/ plus grands que moi le grand frère a quarante-cinq ans. »

(190) *M27A, Korçë, 5 août 2014*

O/	***vaver***	***phral***	*isi*	*saranda*	*brəš-en-g-o.*
ART.DEF.M.SG	autre[M.SG]	frère(M.SG)	être.3.PRES	quarante	année-M.PL.OBL-GEN-M.SG

« Le/ l'autre frère a quarante ans. »

Les deux adjectifs portent un sème spécifique : le référent de *o baro phral*, le grand frère, est contrastif du référent de *o vaver phral*, l'autre frère, mentionné dans la phrase suivante. Il existe donc un focus étroit à valeur contrastive portant sur les adjectifs *baro*, grand, et *vaver*, autre. Dans les exemples (191) et (192), le locuteur arliophone est un homme de 46 ans.

(191) *M46A, Korçë, 24 juillet 2014*

O	***maj***	***puran-e***	***Rom-a***	*kaj*	*si*
ART.DEF.PL	COMP	ancien-PL	Rom-PL	REL	être.3.PRES
andi	*Albanja*	*isi*	*o/*	*o*	*Mečkar-ja.*
dans.ART.DEF.F.SG	Albanie(F.SG)	être.3.PRES	ART.DEF.PL	ART.DEF.PL	Mečkar-M.PL

« Les plus vieux Roms qui sont en Albanie sont les/ les Mečkars […]. »

(192) *M46A, Korçë, 24 juillet 2014*

O	***vaver***	***Rom-a***	*andi*	*Albanja*	*nanaj*
ART.DEF.PL	autre[PL]	Rom-PL	dans.ART.DEF.F.SG	Albanie(F.SG)	être.NEG.3.PRES
po	*but*	*se*	*šelubiš*	*seutrijanda*	*breš.*
COMP	beaucoup	que	cent_vingt	cent_trente	année[M.PL]

« Les autres Roms en Albanie/ ne sont pas là depuis plus de cent vingt cent trente ans. »

L'identification du référent se fait en opérant une restriction pragmatique au sein des référents du nom. Le référent *o maj purane Roma*, les plus anciens Roms, est contrastif du référent *o vaver Roma*, les autres Roms, mentionné deux phrases plus loin. Il existe donc un focus étroit à valeur contrastive portant sur les adjectifs *purane*, anciens, et *vaver*, autre.

Les deux couples d'exemples DEF A N que nous venons de voir sont en opposition de la même façon que le couple DEF N A (181) – (182). Le Tableau 36 montre que les exemples ci-dessus ont un profil

typique de DEF A N en ce qui concerne la structure de l'information (en bleu dans le tableau), mais sont plus proches de DEF N A en ce qui concerne l'interprétation sémantique (en vert dans le tableau).

Le Tableau 37 montre que les quatre exemples ont un profil typique de DEF A N car ils ne font pas partie du focus de la phrase (en bleu dans le tableau), mais plus proches de DEF N A en ce qui concerne la présence d'un focus étroit contrastif (en vert dans le tableau).

Tableau 36 : Complément au Tableau 117, p. 545

Structure	n°	Syntagme	Statut de l'information	Structure de l'information	Interprétation sémantique
DEF A N	(172)	*sa i romani čhib*	donné (A nouveau / accessible)	topique-propos	A – restriction, propriété non-spécifique
	(173)	*to purane breša*	nouveau (A nouveau)	topique-cadre	A – propriété non-spécifique
	(174)	*i tikni čhaj*	accessible (A nouveau)	topique-propos	A – restriction, propriété spécifique
	(175)	*kato baro Del*	nouveau (A nouveau / accessible)	topique-cadre	A – propriété inhérente
	(176)	*o šutalo thud*	donné (A donné)	focus contrastif	A – propriété inhérente
	(177)	*o bilačho manuš*	donné (A donné)	topique-propos	A – restriction, contraste implicite, propriété non-spécifique
	(178)	*o lače manuša*	nouveau (A nouveau)	topique-propos	A – restriction, contraste implicite, propriété non-spécifique
	(179)	*e bare doktorensar*	donné (A nouveau)	focus nouvelle information non- sollicitée	A – propriété inhérente
	(188)	*o jekto kongreso*	nouveau (A nouveau)	topique-propos	A – propriété spécifique inhérente (contraste implicite)
	(189)	*o baro phral*	donné (A donné)	topique-propos	A – restriction, contraste implicite, propriété spécifique
	(190)	*o vaver phral*	donné (A nouveau)	topique-propos	A – restriction, contraste implicite, propriété spécifique
	(191)	*o maj purane Roma*	nouveau (A nouveau)	topique-propos	A – restriction, contraste implicite, propriété spécifique
	(192)	*o vaver Roma*	accessible (A nouveau)	topique-propos	A – restriction, contraste implicite, propriété spécifique
DEF N A	(180)	*o pete bare*	nouveau (A nouveau)	focus nouvelle information non- sollicitée	A – restriction, contraste implicite, propriété spécifique
	(181)	*e čhave bare*	accessible (A donné)	topique-propos	A – restriction, contraste, propriété spécifique
	(182)	*o čhavo cikno*	accessible (A donné)	topique-propos	A – restriction, contraste, propriété spécifique
	(183)	*to amalinja vavera*	nouveau (A nouveau)	focus nouvelle info non sollicitée	A – restriction, propriété spécifique
	(184)	*odole gilja koraxane*	donné (A donné)	focus nouvelle information sollicitée	A – restriction, propriété spécifique
	(185)	*ano foro baro*	nouveau (A nouveau / accessible)	topique-cadre	A – restriction, contraste implicite, propriété spécifique
	(186)	*pire gada romane*	nouveau (A nouveau)	focus nouvelle info non sollicitée	A – ? restriction, contraste, propriété spécifique
	(187)	*to vaćeribe řomano*	nouveau (A accessible)	focus contrastif	A – restriction, contraste, propriété spécifique

Tableau 37 : Complément au Tableau 29, p. 235

Structure	n°	Syntagme	Focus large sur NP		Focus étroit sur A		Pas de focus
			informatif	contrastif	informatif	contrastif	
DEF A N	(172)	*sa i romani čhib*					✓
	(173)	*to purane breša*					✓
	(174)	*i tikni čhaj*					✓
	(175)	*kato baro Del*					✓
	(176)	*o šutalo thud*		✓			
	(177)	*o bilačho manuš*				✓	✓
	(178)	*o lače manuša*				✓	✓
	(179)	*e bare doktorensar*	✓				
	(188)	*o jekto kongreso*				(✓)	✓
	(189)	*o baro phral*				✓	✓
	(190)	*o vaver phral*				✓	✓
	(191)	*o maj purane Roma*				✓	✓
	(192)	*o vaver Roma*				✓	✓
DEF N A	(180)	*o pete bare*	✓			✓	
	(181)	*e čhave bare*				✓	
	(182)	*o čhavo cikno*				✓	
	(183)	*to amalinja vavera*	✓		✓		
	(184)	*odole gilja koraxane*	✓		✓		
	(185)	*ano foro baro*				✓	
	(186)	*pire gada romane*	✓			✓	
	(187)	*to vaćeribe řomano*					

Il existe donc une variation dialectale quant au placement de l'adjectif épithète en Albanie. Une partie des emplois assumés par DEF N A semble assumée par DEF A N chez les Arlis. Cela explique pourquoi l'exemple (174) est un cas de modification indirecte, alors qu'on l'attendrait directe.

Puisqu'on sait que DEF A N est la construction canonique, diachroniquement première, on peut considérer qu'une partie des emplois de DEF A N a été, à un moment donné de l'histoire, assignée à une nouvelle construction, DEF N A, dans les variétés mečkar et čergar I. Elle l'est partiellement en arli, ce qui peut dénoter un changement structurel en cours. La Figure 33 schématise, dans une perspective diachronique, la redistribution des usages de l'adjectif épithète. La langue a innové en créant (ou important d'une langue de contact) un nouvel ordre des constituants.

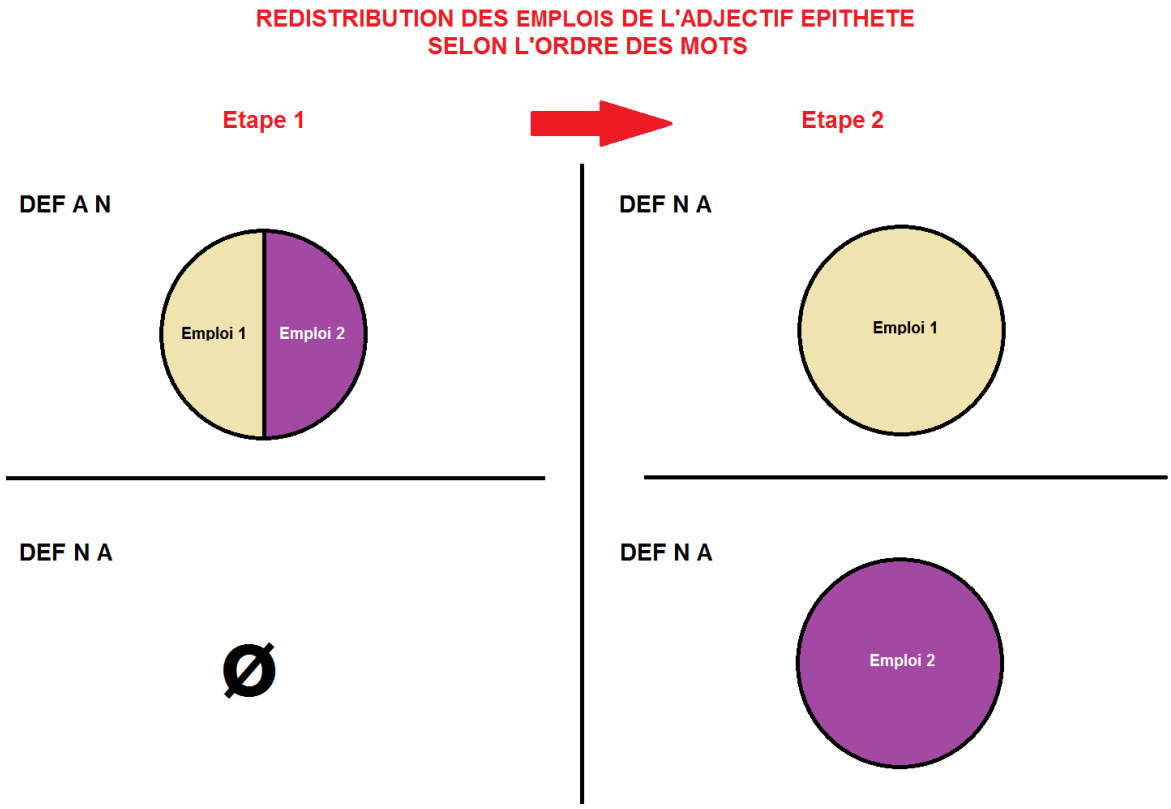


Figure 33 : Redistribution fonctionnelle des fonctions de l'adjectif épithète en diachronie

2.3.5. Pourquoi un deuxième article défini ?

Les deux structures à adjectif postposé DEF N A et DEF N DEF A ont ceci en commun de mettre un focus sur l'adjectif, au contraire des deux structures à adjectif antéposé DEF A N et DEF A DEF N. À quoi sert, alors, l'ajout d'un deuxième article défini ?

Cherchons tout d'abord si l'article défini revêt une fonction différente en syntagme mono- et polydéfini. C'est l'objet du Tableau 38, qui suit la classification de Boretzky (2000b : 44-54) (cf. Tableau 13). On constate que :

- les articles assument quasiment toutes les fonctions possibles.
- on ne trouve que des usages sémantiques, c'est-à-dire *knowledge-dependent*, parmi les syntagmes DEF A DEF N. Au contraire, on trouve des usages à la fois sémantiques et pragmatiques, c'est-à-dire *situation-dependent*, parmi les syntagmes DEF N DEF A.
- on ne trouve que des usages anaphoriques parmi les syntagmes DEF N DEF A à effet de sélection, qu'ils soient dépendants de la connaissance du monde ou de la situation. Il s'agit de

faire référence à une entité qui est, soit déjà mentionnée auparavant, soit « préparée » par la mention d'une entité qui lui est liée.

- on ne trouve que des usages situatifs parmi les syntagmes DEF N DEF A à effet d'emphase, qu'ils soient dépendants de la connaissance du monde ou de la situation. Il s'agit de faire référence à une entité qui est présente dans la situation d'énonciation ou dans la culture générale partagée.
- on trouve une majorité d'usages sémantiques et une majorité d'usages anaphoriques parmi les syntagmes DEF A N.
- on trouve des usages variés parmi les syntagmes DEF N A.
- on trouve dans l'ensemble du corpus une majorité d'usages anaphoriques, quelle que soit la position de l'article défini, plutôt que des usages situatifs.

On peut déduire de la comparaison des syntagmes mono- et polydéfinis que l'antéposition de l'adjectif va de pair avec un usage sémantique ds l'articles définis. Au contraire, la postposition permet des usages variés. L'effet de sélection propre à la polydéfinitude semble aller de pair avec un usage anaphorique de l'article défini, ce qui n'est pas de cas de l'effet d'emphase. Cela pourrait éventuellement nous donner une clé de compréhension des syntagmes à l'interprétation ambiguë : les exemples (166) et (167) ont un usage anaphorique, ce qui nous ferait les interpréter selon un effet emphatique. Les exemples (164), (168) et (169) ont un usage situatif, ce qui nous ferait les interpréter selon un effet de sélection.

Tableau 38 : Usages de l'article défini dans les syntagmes définis

Structure	n°	Syntagme	Usages pragmatiques			Usages sémantiques		
			anaphorique	cataphorique	situatif direct	anaphorique associatif	situatif abstrait	anamnestique
DEF N DEF A	(154)	*o phral o vaver*				✓		
	(155)	*mo phral o xurdo*				✓		
	(156)	*mi phen i/ i tikni*	✓					
	(157)	*mo phral o tikno*	✓					
	(158)	*o phral o baro*				✓		
	(159)	*i čhaj i vaver*	✓					
	(160)	*i škola i bari*	✓					
	(161)	*o phral o tikno*	✓					
	(162)	*o filme o bare*					✓	
	(163)	*i bira i purani*			✓			
	(164)	*o lokali o baro*			✓			
	(165)	*o lokal o baro*	✓					
	(166)	*amaro ćher o baro*	✓					
	(167)	*o čhave o bare*				✓	(✓)	
	(168)	*to ćidipe to/ gadžikane*						✓
	(169)	*o problemi o baro*						✓
DEF A DEF N	(261)	*o tikne o maksemi*				✓		
	(171)	*o barvalo o Skender*						✓
DEF A N	(172)	*sa i romani čhib*	✓			(✓)		
	(173)	*to purane breša*					✓	(✓)
	(174)	*i tikni čhaj*				✓		
	(175)	*kato baro Del*					✓	
	(176)	*o šutalo thud*	✓					

Structure	n°	Syntagme	Usages pragmatiques			Usages sémantiques		
			anaphorique	cataphorique	situatif direct	anaphorique associatif	situatif abstrait	anamnestique
	(177)	*o bilačho manuš*	✓					
	(178)	*o lače manuša*				✓		
	(188)	*o jekto kongreso*	✓					
	(189)	*o baro phral*				✓		
	(190)	*o vaver phral*				✓		
	(191)	*o maj purane Roma*				✓		
	(192)	*o vaver Roma*				✓		
DEF N A	**(180)**	*o pete bare*					✓	
	(181)	*e čhave bare*				✓		
	(182)	*o čhavo cikno*				✓		
	(183)	*to amalinja vavera*					✓	
	(184)	*odole gilja koraxane*	(✓)					
	(185)	*ano foro baro*					✓	
	(186)	*pire gada romane*				(✓)		
	(187)	*to vaćeribe řomano*					✓	

Campos et Stavrou (2004 : 150) décrivent le phénomène suivant pour le grec : *« in polydefinites the adjective is always a restrictive modifier, whereas in monadic constructions it can be restrictive or non-restrictive »*. Un emploi *restrictif* vise à circonscrire le domaine de la référence lorsque plusieurs interprétations sont possibles. Un modifieur non-restrictif s'emploie au contraire pour enrichir la connaissance et la description du référent, en lui attribuant des propriétés, des caractéristiques supplémentaires (Givón 1990 : 10). La fonction du syntagme polydéfini est de restreindre l'ensemble qui constitue la dénotation du nom – plus précisément, grâce à la partie DEF A du syntagme. Au contraire en grec, les syntagmes monodéfinis peuvent être restrictifs ou non.

Comment cela se manifeste-t-il concrètement ? Le contexte d'usage de la polydéfinitude n'est pas le même que celui de la monodéfinitude. Lorsque l'identification du référent est relativement aisée, c'est-à-dire lorsque la dénotation du nom comprend seulement deux entités, un syntagme monodéfini « suffit ». En effet, si une entité a déjà été mentionnée, « l'autre » référent dont il est question ne peut être que la deuxième entité. En revanche, lorsque l'identification est délicate, c'est-à-dire lorsque la dénotation du nom comprend au moins trois entités, un syntagme polydéfini est plus approprié. En effet, « l'autre » peut être plusieurs entités possibles. Le Tableau 39 présente le nombre d'autres entités possibles auxquelles peuvent faire référence les syntagmes du corpus. L'article défini qui accompagne le nom joue donc un rôle d'identification. Cela rappelle la fonction d' « identification » qu'assigne Manolessou (2000) au syntagme polydéfini, et explique pourquoi le syntagme polydéfini donne l'impression d'être plus spécifique qu'un syntagme monodéfini.

Tableau 39 : Degré de restrictivité des syntagmes définis

Structure	n°	Syntagme	Autres référents en contraste
DEF N DEF A	**(154)**	*o phral o vaver*	2 frères (et 6 sœurs)
	(155)	*mo phral o xurdo*	2 frères
	(156)	*mi phen i/ i tikni*	1 soeur et 1 frère
	(157)	*mo phral o tikno*	2 sœurs
	(158)	*o phral o baro*	2 frères
	(159)	*i čhaj i vaver*	3 filles (et 4 fils)
	(160)	*i škola i bari*	3 types d'école : pré-primaire, primaire, secondaire
	(161)	*o phral o tikno*	2 frères
	(162)	*o filme o bare*	–
	(163)	*i bira i purani*	–
	(164)	*o lokali o baro*	de nombreux cafés
	(165)	*o lokal o baro*	de nombreux cafés
	(166)	*amaro ćher o baro*	1 maison : actuelle
	(167)	*o čhave o bare*	1 type d'enfants : plus petits
	(168)	*to ćidipe to/ gadžikane*	1 type de réunion : rom
	(169)	*o problemi o baro*	beaucoup de problèmes
DEF A DEF N	**(261)**	*o tikne o maksemi*	1 type d'enfants
	(171)	*o barvalo o Skender*	[plusieurs interprétations : cf. p.242] 2 frères (et 5 sœurs)
DEF A N	**(172)**	*sa i romani čhib*	–

Structure	n°	Syntagme	Autres référents en contraste
	(173)	*to purane breša*	1 époque : actuelle
	(174)	*i tikni čhaj*	4 autres enfants
	(175)	*kato baro Del*	–
	(176)	*o šutalo thud*	1 autre entité : l'eau
	(177)	*o bilačho manuš*	1 type d'homme : bon
	(178)	*o lače manuša*	1 type d'homme : méchant
	(188)	*o jekto kongreso*	sept congrès suivants
	(189)	*o baro phral*	1 frère : le cadet
	(190)	*o vaver phral*	1 frère : l'aîné
	(191)	*o maj purane Roma*	1 groupe de Roms : ceux arrivés tard
	(192)	*o vaver Roma*	1 groupe de Roms : ceux arrivés tôt
DEF N A	**(180)**	*o pete bare*	1 type de boules : petites
	(181)	*e čhave bare*	1 fils : le petit
	(182)	*o čhavo cikno*	1 fils : le grand
	(183)	*to amalinja vavera*	1 type de gens : les Roms
	(184)	*odole gilja koraxane*	1 type de chanson : romani
	(185)	*ano foro baro*	1 type de ville : petite
	(186)	*pire gada romane*	1 type de vêtements : non-roms
	(187)	*to vaćeribe řomano*	1 langue : albanaise

On constate que lorsque l'adjectif est employé en syntagme polydéfini, il peut contraster avec une pluralité d'éléments – mais pas toujours, ainsi lorsque le syntagme a une fonction d'emphase. Au contraire, les syntagmes monodéfinis ne contrastent jamais avec une pluralité d'éléments : soit ils ne contrastent pas du tout, soit ils contrastent avec un seul élément. Notons l'exception que représentent les exemples (174) et (188), tous deux prononcés par des locuteurs arli. Au moment où le locuteur prononce l'exemple (188), il n'est d'ailleurs pas encore dans une perspective d'énumération donc de contraste : l'adjectif y joue un rôle de quantifieur.

2.3.6. Réponse à la problématique

Nous pouvons affiner les hypothèses proposées des recherches sur le romani (section 2.1, p. 142) :

1) Il s'agit d'un calque du roumain, qui accompagne l'emploi du démonstratif, en cas d'emphase.

Nous ignorons encore si la structure a été empruntée au roumain. Nous n'avons dans le corpus aucun exemple où la construction accompagne l'emploi du démonstratif : nous n'avons au contraire que des emplois avec double article défini, ainsi qu'un emploi du possessif (structure POSS N DEF A dans l'exemple (155) : on ne peut donc affirmer qu'elle accompagne nécessairement le démonstratif.

2) Il s'agit d'un calque du grec, provoqué par la postposition de l'adjectif.

Nous ignorons encore si la structure a été empruntée au grec, s'il s'agit d'une innovation indépendante du romani, ou encore d'une combinaison de ces scénarios : cela fera l'objet de la discussion en sections 1 et 4.2 du chapitre VI, p. 407 et p. 479). Cette structure n'est pas obligatoire en cas de postposition de l'adjectif puisque DEF N A existe en romani d'Albanie (cf. section 2.2.2, p. 231). Nous avons en outre trouvé deux occurrences de structure polydéfinie avec adjectif antéposé, en (261) et (171) : on ne peut donc affirmer que la structure ne se rencontre qu'en cas de postposition de l'adjectif, même si DEF A DEF N semble rarissime dans la langue spontanée.

3) La construction est possible mais non obligatoire, en cas de postposition de l'adjectif et de focus sur cet adjectif.

Cette affirmation a pu être en grande partie confirmée par notre étude.

4) Il s'agit d'un *afterthought* conventionnalisé.
5) Il s'agit d'un adjectif appositionnel traité comme un nom.

Ces affirmations font l'objet de la suite de l'étude.

6) Il s'agit d'un calque du grec, provoqué par la postposition de l'adjectif, en cas d'emphase. L'adjectif est traité comme un nom, mais il ne s'agit pas pour autant d'une apposition lâche ni d'un *afterthought* disjoint.

Nous ignorons encore si la structure a été empruntée au roumain ou au grec, s'il s'agit d'une innovation indépendante du romani, ou encore d'une combinaison de ces scénario.

Nous pouvons également esquisser, ici en bleu, un début de réponse à la problématique (Figure 34).

Questions structurelles

1) Quels éléments peuvent entrer dans un syntagme polydéfini romani ?

La structure polydéfinie est constituée d'un nom précédé d'un article défini (éventuellement précédé d'une préposition) et d'un adjectif précédé d'un article défini (éventuellement précédés d'une préposition). Un déterminant peut venir remplacer l'article défini introduisant le nom : le possessif (POSS N DEF A). Le romani ne se distingue pas ici du grec.

2) Dans quel ordre les trouve-t-on ? Existe-t-il des structures DEF A DEF N ?

Le corpus spontané présente des syntagmes avec adjectif postposé DEF N DEF A, et quelques cas d'ordre DEF A DEF N.

3) L'ordre des mots est-il plus souple en syntagme poly- que monodéfini ? Le romani présente-t-il la même souplesse que le grec ?

Le romani albanais permet l'antéposition de l'adjectif (DEF A N) aussi bien que sa postposition (DEF N A) en syntagme monodéfini. Il se distingue par là du grec, qui ne permet que l'antéposition de l'adjectif (DEF A N, *DEF N A). Le romani est donc plus souple que le grec en syntagme monodéfini.

En syntagme polydéfini, DEF A DEF N comme DEF N DEF A sont possibles en romani. Toutefois, l'antéposition de l'adjectif semble marginale par rapport à sa postposition, car nous avons relevé très peu de cas d'antéposition (DEF A DEF N). Le romani se distingue par là du grec, qui antépose volontiers l'adjectif (DEF A DEF N) autant qu'il le postpose (DEF N DEF A). L'ordre des mots est donc moins souple en syntagme polydéfini que monodéfini en romani – c'est le contraire en grec.

4) Lorsqu'une expression polydéfinie contient plusieurs adjectifs, trouve-t-on les mêmes contraintes d'ordre qu'en grec ?

Le corpus spontané ne présente pas de cas de succession d'adjectifs.

5) L'adjectif d'un syntagme polydéfini peut-il avoir son propre complément ?

Aucun adjectif du corpus spontané n'a son propre complément.

6) L'emplacement postnominal est-il la zone des « options » ? Quel est le degré de nominalité des éléments qui s'y trouvent ?

7) Existe-t-il une restriction quant au type de nom pouvant entrer dans un syntagme polydéfini ?

Nous avons trouvé dans le corpus spontané presque tous les types de noms : un nom propre et des noms communs, un nom indénombrable et des noms dénombrables, un nom d'événement, des noms abstraits, animés, inanimés. Nous n'avons pas trouvé de nom massif ni de nom générique. Ils ont des poids différents et sont aussi bien hérités qu'empruntés.

8) Existe-t-il des contraintes morphologiques ou étymologiques quant au type de l'adjectif ?

Le corpus spontané ne présente que des adjectifs hérités. Nous n'avons pu établir de corrélation entre leur poids et leur position dans le syntagme.

9) Existe-t-il des contraintes syntaxiques ? L'adjectif doit-il être prédicatif ?

Le corpus spontané ne présente que des adjectifs prédicatifs, à l'exception de *vaver*, autre.

10) Existe-t-il des contraintes sémantiques ? Doit-il être restrictif ? Certains adjectifs sont-ils exclus, comme les adjectifs relationnels ou ceux lexicalisés et figés avec le nom ?

Les adjectifs articulés sont orientés vers l'objet et non vers le locuteur, restrictifs et intersectifs (ou intersectifs en apparence subsectifs). Ils permettent une interprétation conjointe. Nous avons toutefois observé plusieurs cas où le corpus polydéfini présente un adjectif non-restrictif, notamment lorsqu'il sert à caractériser, à mettre une étiquette (*epithet*). Dans d'autres cas, on peut parler de restriction seulement si l'on considère un ensemble plus large que celui des seuls référents possibles du nom.

Le corpus spontané présente un cas d'adjectif relationnel et deux cas d'adjectifs semi-lexicalisés avec le nom, ce qui montre qu'ils ne sont pas exclus de la polydéfinitude. Toutefois, dans les deux cas, ils impliquaient un contraste implicite ou explicite.

11) Existe-t-il des contraintes quant au type de déterminant ? Peut-il y avoir répétition d'autre chose que de l'article défini, notamment l'article indéfini ?

Seul l'article défini est répliqué dans notre corpus ; aucun déterminant indéfini ne semble pouvoir être répété. Le possessif peut prendre la place de l'une des occurrences de l'article défini – mais il n'est pas pour autant répliqué. C'est l'article défini qui est alors répété : POSS N DEF A. La structure ne se produit donc qu'en contexte défini : un déterminant défini simple ou complexe en première position et un article défini en seconde position.

12) L'article défini du romani est-il explétif ?

Le corpus spontané montre que l'article défini assume toutes les fonctions présentées par Boretzky (2000b : 44-54). Dans les syntagmes polydéfinis à adjectif antéposé, il revêt un usage sémantique (*knowledge-dependent*) ; dans ceux à adjectif postposé, il revêt aussi bien un usage sémantique que pragmatique (*situation-dependent*).

13) Peut-il y avoir répétition du marquage de cas et de la préposition ?

La polydéfinitude donne lieu à la répétition de la préposition lorsqu'elle a fusionné avec l'article défini. Si un syntagme polydéfini est introduit par une préposition monosyllabique cliticisée à DEF_1, on la trouve cliticisée à DEF_2, donnant lieu à une structure PREP.DEF_1 N PREP.DEF_2 A.

14) La partie adjectivale est-elle un cas de modification directe ou indirecte ?

Nous avons analysé la partie adjectivale des syntagmes polydéfinis comme un cas de modification indirecte.

15) Elle-t-elle un cas d'apposition étroite ?

16) Existe-t-il des syntagmes polydéfinis lâches ? Leur interprétation est-elle non-restrictive ?

17) Y a-t-il un lien entre polydéfinitude et ordre des constituants nominaux ? Lequel ?

Questions sémantiques

1) Existe-t-il une nuance de sens entre syntagme monodéfini et syntagme polydéfini ?

Les syntagmes polydéfinis donnent l'impression d'être plus spécifiques, probablement parce qu'ils sont restrictifs.

2) *Poly*définitude implique-t-elle *sur*définitude ?

La polydéfinitude n'implique pas, à notre sens, de sur-définitude. On pourrait aller jusqu'à dissocier définitude et article défini, en qualifiant celui-ci d'explétif. Il pourrait s'agir d'un marqueur de focalisation.

3) Les syntagmes monodéfinis sont-ils restrictifs, tandis que les syntagmes polydéfinis peuvent être restrictifs ou non-restrictifs ?

En syntagme monodéfini comme en syntagme polydéfini, l'interprétation est soit restrictive soit non-restrictive. Le syntagme polydéfini a pour caractéristique d'effectuer une restriction pragmatique parmi les référents possibles du nom. On peut parler d'une fonction d'identification, qui est la fonction majeure de cette structure. Elle est donc notamment employée lorsque les référents possibles du nom sont plus de deux, et que l'identification est alors considérée comme délicate. Dans les cas (beaucoup moins nombreux) où le syntagme a un effet d'emphase, cette fonction d'identification due à la restriction pragmatique ne semble pas être présente.

4) Si l'adjectif est antéposé au nom, l'interprétation est-elle seulement restrictive ? Si l'adjectif est postposé au nom, l'interprétation peut-elle être restrictive ou non-restrictive ?

En syntagme polydéfini, lorsque l'adjectif est antéposé, l'interprétation est plutôt non-restrictive ; lorsque l'adjectif est postposé, elle est plutôt restrictive. Mais ce ne semble pas être une règle absolue. En syntagme monodéfini, lorsque l'adjectif est antéposé, l'interprétation est restrictive ou non-restrictive ; lorsque l'adjectif est postposé, elle est plutôt restrictive. Là aussi, il ne s'agit que de tendances. C'est donc la position de l'adjectif plutôt que la réplication de l'article défini qui semble jouer un rôle.

Questions fonctionnelles

1) La polydéfinitude est-elle optionnelle en romani comme en grec ? Si oui, qu'est-ce qui provoque le choix du locuteur ?

Le syntagme polydéfini est optionnel en romani comme en grec.

2) S'agit-il d'une variation optionnelle de type *syntactic doubling* (Barbiers 2013 : 7) ?

3) Quelle est la fonction des éléments situés dans l'emplacement postnominal (épithète, apposition, *afterthought*…) ?

Le syntagme polydéfini a pour fonction l'identification. C'est pourquoi il comprend, le plus souvent, un adjectif postposé : la postposition est le lieu du focus, où les sèmes spécifiques sont mis en avant. Nous avons vu qu'un syntagme polydéfini peut avoir d'autres fonctions : précision ou auto-correction, récapitulation, *topic-shift*.

4) Le syntagme polydéfini implique-t-il un contraste ?

Le focus implique souvent un contraste, mais celui-ci n'est pas nécessairement explicite. En effet, on a pu constater que l'effet de sélection de DEF N DEF A va de pair avec un contraste *in presentia* (référents du même ensemble mentionnés en discours ou présents dans la situation d'énonciation) ; l'effet d'emphase de DEF N DEF A va de pair avec un contraste *in absentia*.

5) Quel est le statut de l'information qu'il porte ?

Le statut de l'information du nom n'est pas pertinent, dans la mesure où les syntagmes polydéfinis ne transmettent pas nécessairement une information nouvelle, mais aussi une information donnée ou accessible. L'adjectif, lui, est toujours nouveau lexicalement : il apporte donc une information nouvelle, une propriété qui permet l'identification du référent.

6) Fait-il plutôt partie du topique ou du focus de la phrase ?

Les syntagmes polydéfinis peuvent faire partie soit du topique, soit du focus de la phrase.

7) Implique-t-il un focus sur l'adjectif ? Si oui, de quel type ?

Le syntagme polydéfini implique un focus sur la partie adjectivale du syntagme, soit informatif soit contrastif – sans que l'on ait pu déceler de lien avec la position de l'adjectif.

8) Le focus est-il situé sur le la totalité du syntagme polydéfini ou sur sa partie adjectivale ? Dans le deuxième cas, la partie nominale constitue-t-elle la présupposition ?

Il apparaît que c'est la topologie du syntagme nominal qui détermine quelle sous-partie constitue la présupposition, et quelle partie constitue le focus. La place postnominale est une position de focus, ou rhématique, tandis que la position anténominale est une position non-focale, ou thématique. La première est le lieu de la modification indirecte, tandis que la seconde est le lieu de la modification directe. Du moins est-ce le cas dans les variétés mečkar et čergar I, la variété arli présentant (encore) une distinction moins nette entre les deux. On peut ainsi représenter les quatre structurations possibles d'un syntagme nominal avec adjectif épithète sans rupture interne (Tableau 40).

Tableau 40 : Quatre structurations possibles du syntagme nominal avec adjectif épithète

	Adjectif antéposé	Adjectif postposé
Sans répétition ART.DEF	DEF A N *o baro Del, to purane breša* adjectif en rôle thématique (présupposé) article défini restrictif ou non	DEF N A *o pete bare, o djalekto romano* adjectif en rôle rhématique (focus) article défini restrictif ou non
Avec répétition ART.DEF	DEF A DEF N *o tikne o maksemi, o barvalo o Skender* adjectif en rôle thématique (présupposé) article défini restrictif	DEF N DEF A *mo phral o tikno, o filme o bare* adjectif en rôle rhématique (focus) article défini restrictif

Questions socio-linguistiques

1) Tous les locuteurs albanais du romani emploient-ils la structure polydéfinie ?

2) Observe-t-on une différence de profil socio-linguistique parmi les locuteurs employant cette structure : différence de genre, d'âge ou de niveau d'éducation ?

Le corpus spontané permet quelques premières constatations quant au profil des locuteurs pratiquant la polydéfinitude. Il s'agit aussi bien d'hommes que de femmes, appartenant à toutes les classes d'âge (sauf les plus jeunes ?).

3) Quelles variétés présentent le plus cette structure ?

Dans le corpus spontané, la polydéfinitude est employée par des locuteurs de plusieurs variétés : arli, čergar I, mečkar, arli & čergar I.

4) Les locuteurs qui emploient cette structure peuvent-ils être influencés par la connaissance d'une ou plusieurs langues étrangères, et si oui, lesquelles ?

5) Quel est le profil « type » de la personne employant le plus volontiers la structure polydéfinie ?

6) Quel est le profil « type » de la personne postposant le plus volontiers l'adjectif ?

Dans le corpus spontané, la monodéfinitude avec adjectif postposé est pratiquée par une plus grande diversité de locuteurs, d'âges et de variétés différentes.

7) Quel est le profil « type » de la personne antéposant le plus volontiers l'adjectif ?

Dans le corpus spontané, la monodéfinitude avec adjectif antéposé est pratiquée quasiment exclusivement par les Arlis et les locuteurs de plus de 40 ans.

Questions diachroniques

1) À quand remonte cette structure ? Est-il possible de la dater ?
2) D'où provient-elle ? S'agit-il d'une innovation interne ou d'un « emprunt » ?
3) Quel a été le processus d'acquisition de la structure ?
4) Existe-t-elle dans d'autres variétés du romani ? Constate-t-on des différences ?
5) La syntaxe du syntagme nominal est-elle stable en romani ?
6) La syntaxe du syntagme nominal est-elle stable en romani d'Albanie ?

Questions typologiques

1) La polydéfinitude existe-t-elle dans d'autres langues que le romani et le grec ?
2) Quels sont les différents « types » de polydéfinitude dans les langues d'Europe ?
3) Où se situe le romani dans cette typologie ?
4) Comment nommer au mieux la structure qui fait l'objet de la présente étude ?

Figure 34 : Esquisse de réponse à la problématique

3. Syntagmes avec rupture interne

Cette partie se donne pour objet l'étude des syntagmes présentant une rupture interne : celle-ci est située entre la sous-partie « nominale » et la sous-partie « adjectivale » du syntagme polydéfini (DEF N/ DEF A, DEF A/ DEF N). Les études sur le grec moderne nous apprennent que dans cette langue, divers éléments peuvent venir s'insérer entre les deux sous-parties. La rupture peut être située avant l'adjectif (DEF N/ A) ou avant le nom (DEF A/ N) en syntagme monodéfini. On considère comme une rupture toute pause durant plus de 0,20 seconde. Le Tableau 41 présente la longueur de la pause dans les syntagmes étudiés : la longueur moyenne des pauses est de 0,35 seconde.

Tableau 41 : Pause et insertion dans le corpus d'appositions lâches

Structure	n°	Syntagme	Pause	Insertion
DEF N/ DEF A	**(193)**	*to phral/ o barvalo*	0,27 sec.	non
	(194)	*me phrales/ e xurdes/ e Benes*	0,35 sec. 0,36 sec.	non non
	(195)	*o duj phrala Aurora/ o bare*	0,36 sec.	oui
	(196)	*e čhej kaj si ma i bari/*	non	oui
	(197)	*i čhaj/ bučhola Nəɻtila i baɻi.*	0,24 sec.	oui
DEF N_1/ DEF N_2	**(198)**	*e makseme/ e/ e cekne*	0,29 sec.	non
DEF A/ N	**(201)**	*o baro/ doktori*	0,44 sec.	non
DEF N/ A	**(202)**	*o pištolja/ purane*	0,20 sec.	non
	(203)	*i tradita/ ə ɻomani*	0,64 sec.	oui
	(204)	*i čhib ə/ gadžikani grećist*	0,38 sec.	oui

3.1. Syntagmes polydéfinis

3.1.1. DEF N/ DEF A

Dans le corpus collecté en Albanie, la frontière entre les deux sous-parties du syntagme peut être soit une pause soit un ensemble de mots insérés, soit les deux. Dans le cas d'une insertion, on pourrait également parler d'une disjonction de la sous-partie adjectivale. Le corpus transcrit et traduit, ainsi que le détail de l'analyse de chaque exemple, se trouve en section 1.2.9, p. 599.

3.1.1.1. Cas sans insertion

Dans les exemples (193) et (194), le nom et et le deuxième article défini sont séparés par un silence.

(193) *F44AČ, Korçë, 1er août 2014*

So	*ker-d-as*	*t-o/*	***t-o***	***phral/***	***o***	***barval-o****?*
quoi	faire-PRET-3SG	POSS.2SG-M.SG	POSS.2SG-M.SG	frère(M.SG)	ART.DEF.M.SG	riche-M.SG

« Qu'est-ce qu'il a fait ton/ ton ton frère/ le riche ? » [1]

Nous avons vu en (171) que la locutrice formulait une attaque contre son beau-frère, et qu'elle considère la richesse comme une propriété inhérente de celui-ci. Comme en (171), on peut interpréter l'adjectif comme un cas d'*epithet*, adjectif de caractérisation ou étiquette (cf. section 2.4.4, p. 188) (Chatzikyriakidis 2014 : 1), qui n'est donc pas restrictif. Mais la critique se fait ici plus âpre : comme son interlocutrice a choisi d'ignorer la question, la locutrice se voit contrainte de reposer la question… en faisant tout pour capter son attention. Cela explique selon nous les phénomènes syntaxiques suivants, à l'œuvre en (193) mais pas en (171) :

- L'inversion sujet-verbe : elle vise à mettre en tête de phrase, donc à focaliser, ce qui est déjà le focus de la phrase, *so kerdas*, qu'est-ce qu'il a fait.
- L'utilisation du nom commun *phral*, frère, au lieu du nom propre Skender : il s'agit d'un lien de parenté qui vise à impliquer l'interlocutrice.
- L'utilisation du possessif *to*, ton, au lieu du simple l'article défini *o* : là aussi, la locutrice veut impliquer l'interlocutrice en soulignant la relation entre elle et le référent, objet de la critique.
- L'hésitation dont fait preuve la locutrice pour prononcer cette critique : répétition du possessif (trois occurrences), deux pauses (après le premier possessif et au milieu du syntagme polydéfini). Elle cherche ses mots, réfléchit à la manière la plus appropriée de qualifier le référent : quel nom commun et quel adjectif employer.

Ces éléments font pencher pour un focus assertif plutôt que pour un focus contrastif. Il s'agit d'une modification indirecte.

(194) *F60A, Korçë, 21 juillet 2014*

Barjar-d-om	***m-e***	***phral-es/***	***e***	***xurd-es****/*	*e*
élever-PRET-1SG	POSS1SG- OBL	frère-M.SG.OBL	ART.DEF.OBL	tout_petit-M.SG.OBL	ART.DEF.OBL

[1] L'exemple (193) est issu de la même conversation que l'exemple (171). Le corpus transcrit et traduit se trouve p. 581.

Ben-es	*em*	*e*	*Moza*	*barjar-d-om.*
Beni-M.SG.OBL	et	ART.DEF.OBL	Moza(F.SG)	élever-PRET-1SG

« J'ai élevé mon frère/ le tout petit/ Beni[1] et Moza j'ai élevé. »

On remarque que l'adjectif présente une marque de type nominal *-es*, masculin singulier oblique, alors que la marque adjectivals serait *-e*. On peut penser que *-es* est copié du nom *phrales*, frère.

L'analyse informationnelle et sémantique donne les résultats suivants (Tableau 42) :

Tableau 42 : Analyse informationnelle et sémantique des syntagmes DEF N/ DEF A sans insertion

DEF N/ DEF A	Statut de l'information			Structure de l'information	Interprétation sémantique	Contraste
	référent du syntagme	nom (lexicalement)	adjectif (lexicalement)			
(193)	donné ou nouveau	nouveau	donné ou nouveau	topique-propos	présupposé (N) sélection et focus contrastif ou focus assertif (A) ? caractérisation propriété inhérente	*in absentia*
(194)	donné-inactif	nouveau	nouveau	focus de nouvelle information non sollicitée	intersection restriction pragmatique présupposé (N) sélection et focus étroit (A)	*in presentia*

Le référent de l'expression polydéfinie avec rupture interne est donné ou nouveau en fonction de l'intention de la locutrice en (193). Les termes employés, nom et adjectif, sont plutôt nouveaux. Si le référent de l'expression polydéfinie ainsi que le nom lexical, sont donnés ou accessibles, l'adjectif est souvent nouveau du point de vue lexical. Le syntagme constitue le topique ou le focus de la phrase. L'interprétation de l'expression se fait, comme pour les syntagmes DEF N DEF A sans rupture interne, à l'intersection des ensembles de référents, grâce à une restriction pragmatique. On constate que le contraste se fait à la fois *in presentia* et *in absentia*.

3.1.1.2. Cas avec insertion

Dans les exemples (195) à (197), le nom et le deuxième article défini sont séparés par des mots.

(195) *M46A, Korçë, 24 juillet 2014*

Em/	***o***	***duj***	***phral-a***	***Aurora/***	*o*
et	ART.DEF.PL	deux	frère-M.PL	Aurore	ART.DEF.PL

[1] L'interlocuteur est précisément le « tout petit frère » en question. Mais la locutrice joue le jeu d'une interview journalistique en lui parlant de lui à la troisième personne.

bar-e	*isine/*	*maj*	*džan-gl-e.*
grand-PL	être.3.PRET	COMP	savoir-PART.PASS-PL

« Et/ les deux frères Aurore/ les grands étaient plus expérimentés. »[1]

Dans cet exemple, l'adjectif n'apporte pas de restriction, car le héros est le plus jeune des frères et il n'a que deux frères, qui sont donc par définition les deux grands et plus grands que lui. L'adjectif « grands » ne peut opérer aucune restriction au sein de l'ensemble des deux frères. Toutefois, étant donné que le protagoniste est un petit frère des deux frères en question, on peut penser que le locuteur l'inclut inconsciemment dans l'ensemble des {frères}.

(196) *M58MČ, Voskopojë, 15 juillet 2013*

E	***čhej***	*kaj*	*si*	*ma*	***i***	***baɻ-i/***
ART.DEF.F.SG	fille(F.SG)	REL	être.3.PRES	1SG.OBL	ART.DEF.F.SG	grand-F.SG
si	*saranda*	*jek*	*bərš-en-g.*			
être.3.PRES	quarante	un	an-M.PL.OBL-GEN			

« La fille que j'ai la grande/ a quarante-et-un ans. »[2]

Ici, l'adjectif n'effectue pas à proprement parler une restriction au sein de l'ensemble des filles. La fille en question est la plus grande des enfants, plus grande que les deux fils : cela rappelle la Figure 30 de l'exemple (157), où la sélection se fait plutôt par le nom que par l'adjectif. Il s'agit néanmoins d'un syntagme polydéfini à effet de sélection, où l'adjectif est un modifieur direct. L'adjectif est en quelque sorte restrictif, non au sein de l'ensemble des enfants filles que s'effectue la sélection, mais au sein de l'ensemble des enfants. Comme dans l'exemple précédent, la rupture se manifeste par un silence ainsi que par une insertion, ici d'une proposition relative. Elle est un modifieur supplémentaire qui aide à identifier le référent, à l'instar du nom propre apposé en (195).

(197) *F47M, Llakatund, 13 août 2014*

I	***čhaj/***	*buč-ola*	*Nəɻtila*	***i***	***baɻ-i.***
ART.DEF.F.SG	fille(F.SG)	s'appeler-3SG.PRES	Nertila(F.SG)	ART.DEF.F.SG	grand-F.SG

« La fille s'appelle Nertila la grande. »

L'analyse informationnelle et sémantique donne les résultats suivants (Tableau 43) :

[1] Le corpus transcrit et traduit de cet exemple se trouve p. 565.

[2] Cet exemple est issu du même entretien que les exemples (181) et (182). Le corpus transcrit et traduit se trouve p. 596.

Tableau 43 : Analyse informationnelle et sémantique des syntagmes DEF N/ DEF A avec insertion

DEF N/ DEF A	Statut de l'information			Structure de l'information	Interprétation sémantique	Contraste
	référent du syntagme	nom (lexicalement)	adjectif (lexicalement)			
(195)	accessible inférable	donné	nouveau	topique-propos	intersection ? restriction pragmatique présupposé (N) sélection et focus étroit (A)	*in presentia*
(196)	donné-inactif	donné-inactif	nouveau	topique-propos	intersection ? restriction pragmatique présupposé (N) sélection et focus étroit (A)	*in presentia*
(197)	donné-inactif	donné-inactif	nouveau	topique-propos	intersection restriction pragmatique présupposé (N) sélection et focus étroit (A)	*in presentia*

Le référent de l'expression polydéfinie avec rupture interne est donné ou accessible dans notre corpus. Le nom est toujours donné, tandis que l'adjectif est nouveau. Le syntagme constitue le topique de la phrase. L'interprétation de l'expression se fait, comme pour les syntagmes DEF N DEF A sans rupture interne, à l'intersection des ensembles de référents, grâce à une restriction pragmatique. Dans ces exemples, le contraste se réalise seulement *in presentia*.

3.1.1.3. Les « faux amis » DEF N et DEM N

Les syntagmes avec rupture interne ne semblent pas présenter de différence notable par rapport aux syntagmes sans rupture interne. On peut alors faire l'hypothèse qu'il ne s'agit pas de syntagmes polydéfinis mais de deux syntagmes monodéfinis différents, juxtaposés. La structure serait non pas DEF N/ DEF A, mais DEF N_1/ DEF N_2 où N_2 est un adjectif substantivé. DEF N_2 y constituerait un deuxième syntagme nominal, indépendant : il pourrait par exemple constituer la réponse à une question portant sur la propriété transmise par l'adjectif. Le corpus transcrit et traduit de cette sous-section, ainsi que le détail de l'analyse de chaque exemple, se trouve en section 1.2.10, p. 603.

L'exemple (198) présente une structure DEF N_1/ DEF N_2 où N_2 est un adjectif substantivé.

(198) *M58MČ, Voskpojë, 15 juillet 2013*

Kana	*e*	*dej/*	*ašun-d-el*	***e***	***maksem-e/***
quand	ART.DEF.F.SG.OBL	fille(F.SG)	écouter-PRET-3.PRES	ART.DEF.M.SG.OBL	enfant-M.SG.OBL
e/	***e***	***cekn-e***	*an-o*	*por?*	
ART.DEF.M.SG.OBL	ART.DEF.M.SG.OBL	petit-M.SG.OBL	dans-ART.DEF.M.SG	ventre(M.SG)	

« Quand est-ce que la fille/ entend l'enfant/ le/ le petit dans le ventre ? »

La partie nominale du syntagme, l'enfant, fait partie du contexte, du présupposé. Cependant l'adjectif ne vise pas, selon nous, à sélectionner un petit enfant parmi d'autres. En d'autres termes, le référent de l'expression globale n'est pas à l'intersection de {enfants} et de {entités petites}, car ces deux ensembles sont identiques dans ce contexte. Il s'agit d'un fœtus dans le ventre de la mère, donc, aux yeux du locuteur, d'un « enfant » qui est « petit ». En outre, le référent de « le petit » a un sens générique. On pourrait considérer que l'adjectif transmet une propriété inhérente (« le petit enfant » pour « le fœtus »), mais cela serait contradictoire avec son emplacement postposé, qui fait de lui un modifieur indirect. Pour cette raison, nous considérons qu'il ne s'agit pas d'un syntagme polydéfini, mais de deux syntagmes nominaux parallèles, où le second comprend un adjectif substantivé.

La substantivisation de l'adjectif est-elle tout de même possible avec un adjectif à lecture intersective, dans un syntagme à interprétation intersective ? Oui, mais dans ce cas le syntagme est en général introduit par un démonstratif plutôt que par un article défini.

(199) *M22A, Korçë, 6 août 2014*

Odəja	***bar-i***	*isi*	*prandime.*
DEM.F.SG	grand-F.SG	être.3PRES	marié

« La grande est mariée. »

(200) *F47M, Llakatundˈ, 13 août 2014*

I li	***odja***	***ave***	*si*	*Beɫat.*
FOC	DEM.F.SG	autre	être.3PRES	Berat

« L'autre aussi est à Berat. »[1]

Le référent du syntagme se trouve à l'intersection de la dénotation du nom *fille*, qui est sous-entendu mais présent un peu plus haut dans la conversation, et de la dénotation de l'adjectif *autre*. Le lien avec le nom-antécédent se fait au moyen du démonstratif féminin *odja*. Le focus est en outre exprimé par la particule de focalisation *i li*, aussi – ce qui n'empêche pas l'ensemble du syntagme *i li odja ave*, l'autre aussi, de jouer le rôle de topique-propos de l'énoncé.

L'analyse informationnelle et sémantique donne les résultats suivants (Tableau 44) :

Tableau 44 : Analyse informationnelle et sémantique des syntagmes à adjectif substantivé

DEF A DEM A	Statut de l'information			Structure de l'information	Interprétation sémantique
	référent du syntagme	**nom (lexicalement)**	**adjectif (lexicalement)**		
(198)	donné-inactif	donné-inactif	nouveau	focus de nouvelle information sollicitée	sens général et non spécifique non-intersectif

[1] L'exemple (200) est issu d'un entretien déjà vu pour l'exemple (197). Le corpus transcrit et traduit se trouve p. 604.

(199)	donné-inactif	donné-(in)actif	topique-propos	intersection focus étroit (A) à valeur contrastive
(200)	donné-inactif	nouveau	topique-propos	intersection focus étroit (A) à valeur contrastive

Le référent de l'expression DEF N_1/ DEF N_2 ou DEM N est donné dans ces trois exemples. L'adjectif est indifféremment donné ou nouveau lexicalement. Les deux syntagmes DEM N constituent le topique propos de la phrase, et l'adjectif substantivé y est mis sous focus étroit à valeur contrastive, de même que les polydéfinis à effet de sélection et à interprétation restrictive. L'expression s'interprète de manière intersective – contrairement à DEF N_1/ DEF N_2.

3.1.2. DEF A/ DEF N

Le corpus spontané ne fournit pas d'exemple de structure DEF A/DEF N. On ne peut donc pas discuter la place de l'adjectif dans les syntagmes polydéfinis avec rupture interne.

3.1.3. Synthèse

3.1.3.1. Bilan

Le Tableau 118 et le Tableau 119 présentent la synthèse des analyses effectuées. Comme pour les syntagmes polydéfinis sans rupture interne, le référent de DEF N/ DEF A et des « faux-amis » (DEF N_1/ DEF N_2 et DEM N) est toujours donné (sept occurrences) ou accessible (une occurrence), jamais nouveau. L'adjectif est quant à lui le plus souvent nouveau (six à sept occurrences contre une à deux occurrences d'adjectif donné), là encore à l'instar des syntagmes sans rupture. Les syntagmes font plus souvent partie du topique (cinq occurrences) que du focus (deux occurrences) de la phrase.

Les syntagmes polydéfinis avec rupture provoquent une restriction, un effet de contraste, ainsi dans les exemples (193), (194), (195) et (197). L'adjectif y exprime une propriété spécifique. On a noté une emphase mise sur cette dernière en (193), là encore comme son équivalent sans rupture. Le référent se trouve à l'intersection des référents de la dénotation du nom et de ceux de la dénotation de l'adjectif. C'est également le cas lorsque l'adjectif est substantivé en syntagme DEM N, mais ce n'est pas le cas lorsque le locuteur de (198) effectuait une autocorrection avec DEF N_1/ DEF N_2.

Le contraste se fait à la fois *in presentia* et *in absentia*, à l'instar des structures polydéfinies sans rupture interne – mais le plus souvent in presentia, notamment lorsqu'il y a insertion de mots entre les deux sous-parties du syntagme.

3.1.3.2. Focus

Le Tableau 45 présente les types de focus dans les exemples polydéfinis avec rupture interne. On constate que là non plus, le type de focus présent sur l'adjectif et le type de focus dont fait l'objet l'ensemble du syntagme nominal ne sont pas les mêmes. Tous les exemples présentent un focus étroit sur l'adjectif, des deux types possibles.

Tableau 45 : Le focus en syntagme polydéfini avec rupture interne

Structure	n°	Syntagme	Focus large sur NP		Focus étroit sur A		Pas de focus
			informatif	contrastif	informatif	contrastif	
DEF N/ DEF A	(193)	*to phral/ o barvalo*			✓	(✓)	✓
	(194)	*me phrales/ e xurdes/ e Benes*	✓			✓	
	(195)	*o duj phrala Aurora/ o bare*				✓	✓
	(196)	*e čhej kaj si ma i bari/*			✓	(✓)	✓
	(197)	*i čhaj/ bučhola Nəɻtila i baɻi.*				✓	✓
DEF N_1/ DEF N_2	(198)	*e makseme/ e/ e cekne*	✓				
DEM N	(199)	*odəja bari*				✓	✓
	(200)	*odja ave*				✓	✓

3.2. Syntagmes monodéfinis

3.2.1. DEF A/ N

Le corpus de cet exemple, ainsi que le détail de l'analyse, se trouve en p. 584.

(201) *M46A, Korçë, 22 juillet 2013*

O	***bar-o/***	***doktori***	*vaćer-ela*	*man-ge*	[…]
ART.DEF.M.SG	grand-M.SG	docteur(M.SG)	parler-3SG.PRES	1SG.OBL-DAT	

« Le grand/ docteur me dit […]. »

L'analyse informationnelle et sémantique donne les résultats suivants (Tableau 46) :

Tableau 46 : Analyse informationnelle et sémantique des syntagmes DEF A/ N

DEF A/N	Statut de l'information			Structure de l'information	Interprétation sémantique
	référent du syntagme	nom (lexicalement)	adjectif (lexicalement)		
(201)	donné	donné	donné	topique-propos	propriété du référent qui lui est inhérente propriété spécifique épithète de nature valeur emphatique, non contrastive modification directe

Le référent de l'expression est donné, de même que les termes employés. Le syntagme constitue le topique de la phrase. L'interprétation de l'adjectif est celle d'un modifieur direct, ici à valeur emphatique, non contrastive, quasiment d'épithète de nature.

3.2.2. DEF N/ A

Le corpus dont sont issus les exemples, ainsi que le détail de l'analyse, se trouvent en section 1.2.11, p. 606.

(202) *M71A, Korçë, 28 juillet 2014*

Dikh-as	*sine*	*o*	*bumb-e*	*akhal*	*sar*	*alj-ena/*
voir-1PL	être.3.PRET	ART.DEF.PL	bombe-F.PL	comme_ça	comme	venir.PRET-3PL
o	***pištol-ja/***	***puran-e***	*kaj*	*phab-lea*	*isine*	[…].
ART.DEF.PL	pistolet-M.PL	ancien-PL	REL	brûler-3SG	être.3.PRET	

« On voyait les bombes comme ça qui arrivaient/ les pistolets/ anciens qui tiraient […]. »

(203) *F22M, Llakatund, 13 août 2014*

ə/	*m-i*	*familja/*	*respekton-ela*	***i***	***tradita/***	*ə*	***ɻoman-i***	[…]
3PL	POSS.1SG-F.SG	famille(F.SG)	respecter-3SG.PRES	ART.DEF.F.SG	tradition(F.SG)	euh	rom-F.SG	

« Euh/ ma famille/ respecte la tradition euh/ rom […]. »

(204) *M46A, Korçë, 24 juillet 2014*

Ole	*džan-ena/*	*sa*	*i/*	***i***	***čhib***	***ə/***	***gadžikan-i***	*grećist*
3PL	savoir-3PL.PRES	tout	ART.DEF.F.SG	ART.DEF.F.SG	langue(F.SG)	euh	non_rom-F.SG	en_grec

« Eux ils parlent tous la/ la langue euh/ non Rom grecque. »

La juxtaposition de l'adverbe *grećist*, en grec, dans le rôle d'adjectif, vise à mieux identifier le référent exact de « langue non-rom ». Cette tactique est la même que dans les exemples (185) et (194). Cet exemple est le seul du corpus à présenter une succession d'adjectifs épithètes, avec la structure DEF N/ A A. Les deux adjectifs sont relationnels (*gadžikani grećist*, non-rom grecque), en relation de subordination avec le nom (*čhib*, langue).

L'analyse informationnelle et sémantique donne les résultats suivants (Tableau 47) :

Tableau 47 : Analyse informationnelle et sémantique des syntagmes DEF N/ A

DEF N/ A	Statut de l'information			Structure de l'information	Interprétation sémantique	Contraste
	référent du syntagme	nom (lexicalement)	adjectif (lexicalement)			
(202)	nouveau ou accessible-inférable	nouveau	nouveau	focus de nouvelle information non sollicitée	intersection sélection et focus étroit (A) à valeur contrastive ou assertive modification indirecte	*in absentia*
(203)	accessible-inférable	donné-accessible	nouveau	le focus de nouvelle information sollicitée	intersection sélection et focus étroit (A) à valeur contrastive modification indirecte	*in absentia*
(204)	nouveau	nouveau	nouveau $_1$ + donné-inactif $_2$	focus de nouvelle information sollicitée	intersection sélection et focus étroit (A) à valeur contrastive modification indirecte	*in presentia*

Le référent de l'expression est dans nos exemples donné ou accessible, de même que le nom sur le plan lexical. L'adjectif est en revanche nouveau. Le syntagme constitue le focus de la phrase. Tous ces cas de postposition de l'adjectif s'interprètent comme de la modification indirecte. Le contraste s'opère soit *in presentia*, soit *in absentia* : pas de différence, ici, avec les syntagmes monodéfinis sans rupture interne.

3.2.3. Synthèse

3.2.3.1. Bilan

Le Tableau 120 présente la synthèse des analyses. Comme pour les syntagmes monodéfinis sans rupture interne, le référent de DEF A/ N et de DEF N/ A peut-être donné (une occurrence), accessible

(une occurrence), ou nouveau (deux occurrences). Il est plutôt donné lorsque l'adjectif est antéposé, et nouveau lorsque l'adjectif est postposé. L'adjectif, là aussi, est quasiment toujours nouveau lexicalement (quatre occurrences) : il apporte une information nouvelle, une propriété qui permet l'identification du référent. Le syntagme DEF A/ N fait partie du topique de la phrase, au contraire des syntagmes DEF N/ A, qui font tous partie du focus.

Comme pour les syntagmes monodéfinis sans rupture interne, l'antéposition de l'adjectif est associée à une propriété inhérente au référent. L'adjectif a une fonction emphatique et non contrastive, sans recevoir de focus étroit. Il s'agit d'une modification directe. Au contraire, la postposition de l'adjectif est associée à une restriction, un contraste et une propriété spécifique du référent. L'adjectif a une fonction sélective et reçoit un focus étroit, à valeur contrastive ou assertive. Il s'agit donc d'une modification indirecte.

On a pu constater que le référent des syntagmes DEF N/ A fait l'objet d'un contraste, soit *in presentia* (en majorité), soit *in absentia*.

3.2.3.2. Focus

Le Tableau 48 présente les types de focus dans les exemples monodéfinis avec rupture. On constate que le syntagme à adjectif antéposé ne présente pas de focus étroit, au contraire des syntagmes à adjectif postposé.

Tableau 48 : Ordre des constituants et focus en syntagme monodéfini avec rupture interne

Structure	n°	Syntagme	Focus large sur NP		Focus étroit sur A		Pas de focus
			informatif	contrastif	informatif	contrastif	
DEF A/ N	(201)	*o baro/ doktori*					✓
DEF N/ A	(202)	*o pištolja/ purane*	✓		✓		
	(203)	*i tradita/ ə ˌomani*	✓			✓	
	(204)	*i čhib ə/ gadžikani grećist*	✓			✓	

3.3. Synthèse générale

On remarque que les syntagmes de ce corpus présentent tous un seul adjectif.

3.3.1. Type d'adjectif

• **Poids :** On a vu en section 2.2.2, p. 74, que le poids respectif du nom et de l'adjectif jouerait un rôle : l'élément le plus lourd tend à la postposition. Bien que ce corpus soit trop lacunaire pour que l'on puisse tirer de conclusion fiable, on constate un certain équilibre et la possible absence de lien entre poids des éléments et leur position dans le syntagme nominal (Tableau 49 et Tableau 50).

Tableau 49 : Poids et place de l'adjectif avec rupture interne : quelques exemples du corpus spontané

Adjectifs	DEF A/ N	DEF N/ A	DEF N/ DEF A
2 syllabes	*o baro/ doktori* le grand/docteur		e čhej kaj si ma i bari *la fille que j'ai la grande*
3 syllabes		*o pištolja/ purane* les pistolets/ anciens	*to phral/ o barvalo* ton frère/ le riche
4 syllabes		*i čhib ə/ gadžikani grećist* la langue euh/ non Rom grecque	

Tableau 50 : Poids et place du nom avec rupture interne : quelques exemples du corpus spontané

Noms	DEF A/ N	DEF N/ A	DEF N/ DEF A
1 syllabe		*i čhib ə/ gadžikani grećist* la langue euh/ non Rom grecque	*to phral/ o barvalo* ton frère/ le riche
2 syllabes			*o duj phrala Aurora/ o bare* les deux frères Aurore/ les grands
3 syllabes	*o baro/ doktori* le grand/ docteur		

• **Origine :** du point de vue étymologique, les adjectifs du corpus sont tous hérités du fonds indien. Les noms sont aussi bien hérités (*čhaj*, fille) qu'empruntés (*tradita*, tradition).

• **Syntaxe** : dans ce corpus, tous les adjectifs sont prédicatifs.

• **Sémantique** : Sur les cinq cas de syntagme polydéfini lâche, seuls deux sont incontestablement restrictifs, (194) et (197). L'adjectif en (193), « riche », n'est pas restrictif si on l'interprète comme une caractéristique (*epithet*) inhérente du référent en question. Les adjectifs en (195), « grands », et (196), « grande » ne sont restrictifs que si l'on prend en compte un ensemble de référents plus large que celui que dénote exactement le nom.

3.3.2. Modification directe et indirecte

La modification se répartit dans les syntagmes polydéfinis avec rupture interne comme indiqué dans le Tableau 51. Ils ne se distinguent pas des syntagmes polydéfinis du Tableau 33.

Tableau 51 : Modification directe et indirecte en syntagme polydéfini avec rupture interne

Structure	n°	Syntagme	Modification directe	Modification indirecte
DEF N/ DEF A	(193)	*to phral/ o barvalo*		✓
	(194)	*me phrales/ e xurdes/ e Benes*		✓
	(195)	*o duj phrala Aurora/ o bare*		✓
	(196)	*e čhej kaj si ma i bari/*		✓
	(197)	*i čhaj/ bučhola Nəɻtila i baɻi.*		✓

La modification se répartit dans les syntagmes monodéfinis avec rupture interne comme indiqué dans le Tableau 51. Ils ne se distinguent pas des syntagmes monodéfinis du Tableau 32. La position antéposée est celle de la modification directe, tandis que la position postposée est celle de la modification indirecte (Tableau 52).

Tableau 52 : Modification directe et indirecte en syntagme monodéfini avec rupture interne

Structure	n°	Syntagme	Modification directe	Modification indirecte
DEF A/ N	(201)	*o baro/ doktori*	✓	
DEF N/ A	(202)	*o pištolja/ purane*		✓
	(203)	*i tradita/ ə ɻomani*		✓
	(204)	*i čhib ə/ gadžikani grećist*		✓

3.3.3. Place de l'adjectif et variation dialectale

Remarquons tout d'abord que ce corpus présente un cas de succession d'adjectifs épithètes, l'exemple (204). Sa structure est DEF N/ A A, avec deux adjectifs relationnels (*gadžikani grećist*, non-rom grecque) en relation de subordination avec le nom (*čhib*, langue).

Le Tableau 53 présente le profil des personnes ayant prononcé les exemples de cette section.

Tableau 53 : Profil des locuteurs des syntagmes polydéfinis avec rupture intene

Structure	n°	Syntagme	Locuteur	Genre	Age[1]	Variété
DEF N/ DEF A	(193)	*to phral/ o barvalo*	F44AČ	F	44	čergar I & arli
	(194)	*me phrales/ e xurdes/ e Benes*	F60A	F	60	arli
	(195)	*o duj phrala Aurora/ o bare*	M46A	M	46	arli
	(196)	*e čhej kaj si ma i bari/*	M58MČ	M	58	mečkar & čergar I
	(197)	*i čhaj/ bučhola Nəɻtila i baɻi.*	F47M	F	47	mečkar

[1] A l'époque de l'enregistrement (2013 ou 2014).

Comme pour les syntagmes sans rupture interne, on constate pour ce corpus que la polydéfinitude avec rupture est pratiquée :

- aussi bien par les femmes (3 personnes) que par les hommes (2 personnes)
- par des personnes de tous âges (de 16 à 60 ans).
- par des locuteurs de toutes les variétés : arli, arli & čergar I, mečkar, mečkar & čergar I.

Le Tableau 54 présente le profil des personnes ayant prononcé les syntagmes à adjectif substantivé.

Tableau 54 : Profil des locuteurs ayant prononcé les syntagmes à adjectif substantivé

Structure	n°	Syntagme	Locuteur	Genre	Age[1]	Variété
DEF N_1/ DEF N_2	**(198)**	*e makseme/ e/ e cekne*	M58MČ	M	58	mečkar & čergar I
DEM N	**(199)**	*odəja bari*	M22A	M	22	arli
	(200)	*odja ave*	F47M	F	47	mečkar

On constate que dans ce corpus la substantivisation de l'adjectif est pratiquée par des locuteurs de différentes variétés (arli, mečkar, mečkar & čergar I) et d'âge différents (22 à 58 ans).

Le Tableau 55 présente le profil des personnes ayant prononcé les exemples monodéfinis.

Tableau 55 : Profil des locuteurs des syntagmes monodéfinis avec rupture interne ou ellipse

Structure	n°	Syntagme	Locuteur	Genre	Age[1]	Variété
DEF A/ N	**(201)**	*o baro/ doktori*	M46A	M	46	arli
DEF N/ A	**(202)**	*o pištolja/ purane*	M71A	M	71	arli
	(203)	*i tradita/ ə ɻomani*	F22M	F	22	mečkar
	(204)	*i čhib ə/ gadžikani grećist*	M46A	M	46	arli

On constate pour ce corpus que :

- Le syntagme DEF A/ N est prononcé par un locuteur arli d'âge moyen.
- Les syntagmes DEF N/ A sont prononcés par des locuteurs de différentes variétés (arli et mečkar) et d'âge différents (22 à 71 ans).

Là encore, on retrouve les mêmes tendances que pour les syntagmes sans rupture interne.

3.3.4. Rôle de la rupture et de l'insertion

Dans notre corpus de romani d'Albanie, l'apposition lâche présente soit un changement intonatif, soit l'insertion d'un élément entre la partie nominale et la partie adjectivale, mais pas les deux à la fois.

Cette insertion peut être un mot, plusieurs mots, voire tout le reste de l'énoncé – diversité qui rappelle le grec (Lekakou et Szendrői 2007 : 131-132). L'exemple (196) montre que le syntagme peut également ne pas présenter de pause. Nous n'avons pas constaté de changement sémantique ou fonctionnel qui soit significatif entre les syntagmes sans rupture et les syntagmes avec rupture interne – qu'ils soient d'ailleurs polydéfinis ou monodéfinis : ni pour le statut, ni pour la structure de l'information, ni pour l'interprétation sémantique, ni pour la signification de l'ordre des constituants.

Le Tableau 56 présente le nombre d'autres entités possibles auxquelles peuvent faire référence les syntagmes. On constate que, là aussi, les syntagmes polydéfinis contrastent de préférence avec une pluralité d'éléments, tandis que les syntagmes monodéfinis contrastent avec un seul élément voire ne contrastent pas du tout. Les syntagmes à adjectif substantivé se comportent comme les syntagmes monodéfinis.

Tableau 56 : Degré de restrictivité des syntagmes définis avec rupture interne

Structure	n°	Syntagme	Autres référents en contraste
DEF N/ DEF A	**(193)**	*to phral/ o barvalo*	2 frères (et 5 soeurs)
	(194)	*me phrales/ e xurdes/ e Benes*	2 frères (et 5 soeurs)
	(195)	*o duj phrala Aurora/ o bare*	1 frère plus petit
	(196)	*e čhej kaj si ma i bari/*	2 fils
	(197)	*i čhaj/ bučhola Nəɻtila i baɻi.*	2 autres enfants (1 fille, 1 fils)
DEF N_1/ DEF N_2	**(198)**	*e makseme/ e/ e cekne*	–
DEM N	**(199)**	*odəja bari*	1 autre sœur
	(200)	*odja ave*	1 autre fille
DEF A/ N	**(201)**	*o baro/ doktori*	–
DEF N/ A	**(202)**	*o pištolja/ purane*	– (1 type de pistolets : modernes)
	(203)	*i tradita ə/ ɻomani*	1 type de tradition : non-rom
	(204)	*i čhib ə/ gadžikani grećist*	*1er adjectif :* 1 langue : rom *2e adjectif :* 1 langue : albanais

Si les syntagmes avec rupture ne présentent pas de différence sémantique ou fonctionnelle par rapport à ceux sans rupture, alors à quoi sert celle-ci ? S'agit-il d'une simple hésitation quant à la manière de qualifier le référent (nom) ou sa propriété (adjectif) – auquel cas il n'y aurait pas de différence de nature entre DEF N DEF A et DEF N/ DEF A ? Nous laissons ces questions pour une étude ultérieure.

3.3.5. Réponse à la problématique

Nous pouvons poursuivre, ici en vert, notre réponse à la problématique (Figure 35).

Questions structurelles

1) Quels éléments peuvent entrer dans un syntagme polydéfini romani ?

La structure polydéfinie est constituée d'un nom précédé d'un article défini (éventuellement précédé d'une préposition) et d'un adjectif précédé d'un article défini (éventuellement précédés d'une préposition). Un déterminant peut venir remplacer l'article défini introduisant le nom : le possessif (POSS N DEF A). Le romani ne se distingue pas ici du grec.

Le corpus avec rupture interne confirme qu'il est possible d'avoir un possessif à la place du premier article défini (POSS N DEF A), mais également qu'il est possible que deux déterminants introduisent le nom, article défini et quantifieur qui lui succède (DEF QUANT N/ DEF A).

2) Dans quel ordre les trouve-t-on ? Existe-t-il des structures DEF A DEF N ?

Le corpus spontané présente des syntagmes avec adjectif postposé DEF N DEF A, et quelques cas d'ordre DEF A DEF N. Le corpus avec rupture interne ne présente aucun cas de syntagme polydéfini avec adjectif antéposé.

3) L'ordre des mots est-il plus souple en syntagme poly- que monodéfini ? Le romani présente-t-il la même souplesse que le grec ?

Le romani albanais permet l'antéposition de l'adjectif (DEF A N) aussi bien que sa postposition (DEF N A) en syntagme monodéfini. Il se distingue par là du grec, qui ne permet que l'antéposition de l'adjectif (DEF A N, *DEF N A). Le romani est donc plus souple que le grec en syntagme monodéfini.

En syntagme polydéfini, DEF A DEF N comme DEF N DEF A sont possibles en romani. Toutefois, l'antéposition de l'adjectif semble marginale par rapport à sa postposition, car nous avons relevé très peu de cas d'antéposition (DEF A DEF N). Le romani se distingue par là du grec, qui antépose volontiers l'adjectif (DEF A DEF N) autant qu'il le postpose (DEF N DEF A). L'ordre des mots est donc moins souple en syntagme polydéfini que monodéfini en romani – c'est le contraire en grec.

4) Lorsqu'une expression polydéfinie contient plusieurs adjectifs, trouve-t-on les mêmes contraintes d'ordre qu'en grec ?

Le corpus spontané ne présente pas de cas de succession d'adjectifs. Tous les syntagmes polydéfinis avec rupture présentent également un seul adjectif.

5) L'adjectif d'un syntagme polydéfini peut-il avoir son propre complément ?

Aucun adjectif du corpus spontané n'a son propre complément. Les syntagmes avec rupture n'en présentent pas non plus.

6) L'emplacement postnominal est-il la zone des « options » ? Quel est le degré de nominalité des éléments qui s'y trouvent ?

7) Existe-t-il une restriction quant au type de nom pouvant entrer dans un syntagme polydéfini ?

Nous avons trouvé dans le corpus spontané presque tous les types de noms : un nom propre et des noms communs, un nom indénombrable et des noms dénombrables, un nom d'événement, des noms abstraits, animés, inanimés. Nous n'avons pas trouvé de nom massif ni de nom générique. Ils ont des poids différents et sont aussi bien hérités qu'empruntés.

Les syntagmes avec rupture ne présentent que des noms communs, dénombrables et animés spécifiques. Ils sont de poids léger et d'origine héritée.

8) Existe-t-il des contraintes morphologiques ou étymologiques quant au type de l'adjectif ?

Le corpus spontané ne présente que des adjectifs hérités. Nous n'avons pu établir de corrélation entre leur poids et leur position dans le syntagme. Les adjectifs qualificatifs des syntagmes avec rupture sont tous hérités.

9) Existe-t-il des contraintes syntaxiques ? L'adjectif doit-il être prédicatif ?

Le corpus spontané ne présente que des adjectifs prédicatifs, à l'exception de *vaver*, autre.

Les adjectifs qualificatifs des syntagmes avec rupture sont tous prédicatifs.

10) Existe-t-il des contraintes sémantiques ? Doit-il être restrictif ? Certains adjectifs sont-ils exclus, comme les adjectifs relationnels ou ceux lexicalisés et figés avec le nom ?

Les adjectifs articulés sont orientés vers l'objet et non vers le locuteur, restrictifs et intersectifs (ou intersectifs en apparence subsectifs). Ils permettent une interprétation conjointe. Nous avons toutefois observé plusieurs cas où le corpus polydéfini présente un adjectif non-restrictif, notamment lorsqu'il sert à caractériser, à mettre une étiquette (*epithet*). Dans d'autres cas, on peut parler de restriction seulement si l'on considère un ensemble plus large que celui des seuls référents possibles du nom.

Le corpus spontané présente un cas d'adjectif relationnel et deux cas d'adjectifs semi-lexicalisés avec le nom, ce qui montre qu'ils ne sont pas exclus de la polydéfinitude. Toutefois, dans les deux cas, ils impliquaient un contraste implicite ou explicite.

Les syntagmes avec rupture présentent des adjectifs qualificatifs permettant une interprétation intersective conjointe, tous orientés vers l'objet. Plusieurs sont restrictifs, mais d'autres sont ambigus

quant à une interprétation restrictive ou non : un cas de l'adjectif de caractérisation et deux cas de restriction dans un ensemble plus large que les seuls référents possibles du nom.

11) Existe-t-il des contraintes quant au type de déterminant ? Peut-il y avoir répétition d'autre chose que de l'article défini, notamment l'article indéfini ?

Seul l'article défini est répliqué dans notre corpus ; aucun déterminant indéfini ne semble pouvoir être répété. Le possessif peut prendre la place de l'une des occurrences de l'article défini – mais il n'est pas pour autant répliqué. C'est l'article défini qui est alors répété : POSS N DEF A. La structure ne se produit donc qu'en contexte défini : un déterminant défini simple ou complexe en première position et un article défini en seconde position.

Les syntagmes à rupture interne nous amènent à nuancer cette dernière affirmation, puisque l'on peut trouver en première position un déterminant défini associé à un déterminant indéfini, le quantifieur : structure DEF QUANT N/ DEF A.

12) L'article défini du romani est-il explétif ?

Le corpus spontané montre que l'article défini assume toutes les fonctions présentées par Boretzky (2000b : 44-54). Dans les syntagmes polydéfinis à adjectif antéposé, il revêt un usage sémantique (*knowledge-dependent*) ; dans ceux à adjectif postposé, il revêt aussi bien un usage sémantique que pragmatique (*situation-dependent*).

13) Peut-il y avoir répétition du marquage de cas et de la préposition ?

La polydéfinitude donne lieu à la répétition de la préposition lorsqu'elle a fusionné avec l'article défini. Si un syntagme polydéfini est introduit par une préposition monosyllabique cliticisée à DEF_1, on la trouve cliticisée à DEF_2, donnant lieu à une structure PREP.DEF_1 N PREP.DEF_2 A.

Le corpus avec rupture interne montre un cas de réplication des marques flexionnelles : l'un des exemples présente une marque de flexion sur le nom et son déterminant, qui est répétée sur l'adjectif et son déterminant postposés. On en déduit que lorsque le syntagme porte une marque de cas, celle-ci est nécessairement répétée, du moins lorsque l'adjectif est postposé.

14) La partie adjectivale est-elle un cas de modification directe ou indirecte ?

Nous avons analysé la partie adjectivale des syntagmes polydéfinis comme un cas de modification indirecte.

La différence entre modification directe et indirecte est d'ailleurs apparue également dans les syntagmes avec rupture et nous avons analysé la partie adjectivale des syntagmes polydéfinis comme de la modification indirecte.

15) Elle-t-elle un cas d'apposition étroite ?

Si l'on conçoit la différence entre apposition étroite et apposition lâche comme une différence structurelle et non sémantique, alors il existe la même distinction en romani. L'apposition étroite, dont ressortent les syntagmes polydéfinis sans rupture, constitue une seule unité prosodique, portant un seul accent. L'apposition lâche, dont ressortent les syntagmes polydéfinis avec rupture, comprend deux sous-unités prosodiques et porte deux accents nucléaires. Le syntagme polydéfini est une apposition étroite, comprenant une partie nominale et une partie adjectivale.

En revanche, il existe une différence importante entre les syntagmes polydéfinis lâches du romani et ceux du grec moderne. En romani ils peuvent être restrictifs (à l'instar des polydéfinis étroits) alors que ceux du grec ne sont pas restrictifs, si l'on en croit la littérature. Il ne s'agit donc pas, en réalité, du même type d'apposition. Plusieurs de nos exemples constituent un autre type encore que celui des appositions lâches de la littérature : restrictifs comme les polydéfinis classiques (apposition étroite), ils comprennent deux sous-unités prosodiques distinctes comme les polydéfinis lâches (apposition lâche). Pour autant, il ne s'agit pas de deux syntagmes juxtaposés DEF N DEF N où l'adjectif est substantivé. En effet, lorsque l'adjectif substantivé est introduit par un article défini, il n'est alors pas restrictif. Lorsque l'adjectif substantivé est restrictif, il n'est alors pas introduit par un article défini, mais par un démonstratif.

16) Existe-t-il des syntagmes polydéfinis lâches ? Leur interprétation est-elle non-restrictive ?

Il existe des syntagmes polydéfinis lâches en romani comme en grec. Ils présentent une rupture entre la sous-partie « nominale » et la sous-partie « adjectivale » : en général une pause, avec parfois un ensemble de mots insérés. Il peut s'agir d'un ou de plusieurs mots : nom propre, interjection, proposition relative, voire le reste de l'énoncé. Cette diversité rappelle celle du grec. L'un de nos exemples montre que le syntagme peut également ne pas présenter de pause.

Dans notre corpus, l'interprétation des syntagmes polydéfinis avec rupture interne est moins systématiquement restrictive que celle des syntagmes polydéfinis sans rupture interne. Il y a donc une différence avec le grec, où un syntagme avec pause ou insertion implique que l'adjectif apporte une information non essentielle et non-restrictive.

17) Y a-t-il un lien entre polydéfinitude et ordre des constituants nominaux ? Lequel ?

Questions sémantiques

1) Existe-t-il une nuance de sens entre syntagme monodéfini et syntagme polydéfini ?

Les syntagmes polydéfinis donnent l'impression d'être plus spécifiques, probablement parce qu'ils sont restrictifs. Les exemples de syntagmes polydéfinis avec rupture fournis par notre corpus sont en revanche tout aussi spécifiques que les syntagmes monodéfinis avec rupture.

2) *Poly*définitude implique-t-elle *sur*définitude ?

La polydéfinitude n'implique pas, à notre sens, de sur-définitude. On pourrait aller jusqu'à dissocier définitude et article défini, en qualifiant celui-ci d'explétif. Il pourrait s'agir d'un marqueur de focalisation. Les syntagmes avec rupture ne se différencient pas des syntagmes sans rupture.

3) Les syntagmes monodéfinis sont-ils restrictifs, tandis que les syntagmes polydéfinis peuvent être restrictifs ou non-restrictifs ?

En syntagme monodéfini comme en syntagme polydéfini, l'interprétation est soit restrictive soit non-restrictive. Le syntagme polydéfini a pour caractéristique d'effectuer une restriction pragmatique parmi les référents possibles du nom. On peut parler d'une fonction d'identification, qui est la fonction majeure de cette structure. Elle est donc notamment employée lorsque les référents possibles du nom sont plus de deux, et que l'identification est alors considérée comme délicate. Dans les cas (beaucoup moins nombreux) où le syntagme a un effet d'emphase, cette fonction d'identification due à la restriction pragmatique ne semble pas être présente.

Les syntagmes avec rupture ne se différencient pas des syntagmes sans rupture : l'interprétation est soit restrictive soit non-restrictive, que le syntagme soit monodéfini ou polydéfini.

4) Si l'adjectif est antéposé au nom, l'interprétation est-elle seulement restrictive ? Si l'adjectif est postposé au nom, l'interprétation peut-elle être restrictive ou non-restrictive ?

En syntagme polydéfini, lorsque l'adjectif est antéposé, l'interprétation est plutôt non-restrictive ; lorsque l'adjectif est postposé, elle est plutôt restrictive. Mais ce ne semble pas être une règle absolue. En syntagme monodéfini, lorsque l'adjectif est antéposé, l'interprétation est restrictive ou non-restrictive ; lorsque l'adjectif est postposé, elle est plutôt restrictive. Là aussi, il ne s'agit que de tendances. C'est donc la position de l'adjectif plutôt que la réplication de l'article défini qui semble jouer un rôle.

Les syntagmes polydéfinis avec rupture présentent tous un adjectif postposé, et une interprétation soit restrictive, soit non-restrictive. Il est donc difficile d'en déduire une tendance. En syntagme monodéfini, lorsque l'adjectif est antéposé, l'interprétation non-restrictive ; lorsque l'adjectif est postposé, elle est restrictive. La taille du corpus est très restreinte, mais elle confirme l'importance de l'ordre des constituants.

Questions fonctionnelles

1) La polydéfinitude est-elle optionnelle en romani comme en grec ? Si oui, qu'est-ce qui provoque le choix du locuteur ?

Le syntagme polydéfini est optionnel en romani comme en grec.

2) S'agit-il d'une variation optionnelle de type *syntactic doubling* (Barbiers 2013 : 7) ?

3) Quelle est la fonction des éléments situés dans l'emplacement postnominal (épithète, apposition, *afterthought*…) ?

Le syntagme polydéfini a pour fonction l'identification. C'est pourquoi il comprend, le plus souvent, un adjectif postposé : la postposition est le lieu du focus, où les sèmes spécifiques sont mis en avant. Nous avons vu qu'un syntagme polydéfini peut avoir d'autres fonctions : précision ou auto-correction, récapitulation, *topic-shift*.

Les syntagmes avec rupture ne se différencient pas des syntagmes sans rupture.

4) Le syntagme polydéfini implique-t-il un contraste ?

Le focus implique souvent un contraste, mais celui-ci n'est pas nécessairement explicite. En effet, on a pu constater que l'effet de sélection de DEF N DEF A va de pair avec un contraste *in presentia* (référents du même ensemble mentionnés en discours ou présents dans la situation d'énonciation) ; l'effet d'emphase de DEF N DEF A va de pair avec un contraste *in absentia*.

On remarque que le contraste *in presentia*, qui correspond à l'effet de sélection explicite, se trouve plus dans les syntagmes avec rupture que sans rupture.

5) Quel est le statut de l'information qu'il porte ?

Le statut de l'information du nom n'est pas pertinent, dans la mesure où les syntagmes polydéfinis ne transmettent pas nécessairement une information nouvelle, mais aussi une information donnée ou accessible. L'adjectif, lui, est toujours nouveau lexicalement : il apporte donc une information nouvelle, une propriété qui permet l'identification du référent.

Les syntagmes avec rupture ne se différencient pas des syntagmes sans rupture.

6) Fait-il plutôt partie du topique ou du focus de la phrase ?

Les syntagmes polydéfinis peuvent faire partie soit du topique, soit du focus de la phrase.

Les syntagmes avec rupture ne se différencient pas des syntagmes sans rupture.

7) Implique-t-il un focus sur l'adjectif ? Si oui, de quel type ?

Le syntagme polydéfini implique un focus sur la partie adjectivale du syntagme, soit informatif soit contrastif – sans que l'on ait pu déceler de lien avec la position de l'adjectif.

Les syntagmes avec rupture ne se différencient pas des syntagmes sans rupture.

8) Le focus est-il situé sur le la totalité du syntagme polydéfini ou sur sa partie adjectivale ? Dans le deuxième cas, la partie nominale constitue-t-elle la présupposition ?

Il apparaît que c'est la topologie du syntagme nominal qui détermine quelle sous-partie constitue la présupposition, et quelle partie constitue le focus. La place postnominale est une position de focus, ou rhématique, tandis que la position anténominale est une position non-focale, ou thématique. La première est le lieu de la modification indirecte, tandis que la seconde est le lieu de la modification directe. Du moins est-ce le cas dans les variétés mečkar et čergar I, la variété arli présentant (encore) une distinction moins nette entre les deux. On peut ainsi représenter les quatre structurations possibles d'un syntagme nominal avec adjectif épithète sans rupture interne (Tableau 40). Les syntagmes avec rupture ne se différencient pas des syntagmes sans rupture.

Tableau 40 : Quatre structurations possibles du syntagme nominal avec adjectif épithète

	Adjectif antéposé	Adjectif postposé
Sans répétition ART.DEF	DEF A N *o baro Del, to purane breša* adjectif en rôle thématique (présupposé) article défini restrictif ou non	DEF N A *o pete bare, o djalekto romano* adjectif en rôle rhématique (focus) article défini restrictif ou non
Avec répétition ART.DEF	DEF A DEF N *o tikne o maksemi, o barvalo o Skender* adjectif en rôle thématique (présupposé) article défini restrictif	DEF N DEF A *mo phral o tikno, o filme o bare* adjectif en rôle rhématique (focus) article défini restrictif

Questions socio-linguistiques

1) Tous les locuteurs albanais du romani emploient-ils la structure polydéfinie ?

2) Observe-t-on une différence de profil socio-linguistique parmi les locuteurs employant cette structure : différence de genre, d'âge ou de niveau d'éducation ?

Le corpus spontané permet quelques premières constatations quant au profil des locuteurs pratiquant la polydéfinitude. Il s'agit aussi bien d'hommes que de femmes, appartenant à toutes les classes d'âge (sauf les plus jeunes ?).

Le corpus avec rupture confirme que la polydéfinitude est pratiquée par tous les locuteurs du romani d'Albanie : homme et femmes de toutes classes d'âge.

3) Quelles variétés présentent le plus cette structure ?

Dans le corpus spontané, la polydéfinitude est employée par des locuteurs de plusieurs variétés : arli, čergar I, mečkar, arli & čergar I.

Dans le corpus spontané avec rupture interne, la polydéfinitude est employée par des locuteurs de plusieurs variétés : arli, mečkar, čergar I & arli, mečkar & čergar I.

4) Les locuteurs qui emploient cette structure peuvent-ils être influencés par la connaissance d'une ou plusieurs langues étrangères, et si oui, lesquelles ?
5) Quel est le profil « type » de la personne employant le plus volontiers la structure polydéfinie ?

6) Quel est le profil « type » de la personne postposant le plus volontiers l'adjectif ?

Dans le corpus spontané, la monodéfinitude avec adjectif postposé est pratiquée par une plus grande diversité de locuteurs, d'âges et de variétés différentes.

7) Quel est le profil « type » de la personne antéposant le plus volontiers l'adjectif ?

Dans le corpus spontané, la monodéfinitude avec adjectif antéposé est pratiquée quasiment exclusivement par les Arlis et les locuteurs de plus de 40 ans.

Dans le corpus avec rupture interne, la monodéfinitude avec adjectif antéposé est le fait d'un profil beaucoup plus restreint de locuteurs que celle avec adjectif postposé : locuteurs arli d'âge moyen ou avancé dans le premier cas, locuteurs de toutes variétés et de tous âges dans le deuxième cas.

Questions diachroniques

1) À quand remonte cette structure ? Est-il possible de la dater ?
2) D'où provient-elle ? S'agit-il d'une innovation interne ou d'un « emprunt » ?
3) Quel a été le processus d'acquisition de la structure ?
4) Existe-t-elle dans d'autres variétés du romani ? Constate-t-on des différences ?
5) La syntaxe du syntagme nominal est-elle stable en romani ?
6) La syntaxe du syntagme nominal est-elle stable en romani d'Albanie ?

Questions typologiques

1) La polydéfinitude existe-t-elle dans d'autres langues que le romani et le grec ?
2) Quels sont les différents « types » de polydéfinitude dans les langues d'Europe ?
3) Où se situe le romani dans cette typologie ?
4) Comment nommer au mieux la structure qui fait l'objet de la présente étude ?

Figure 35 : Début de la réponse à la problématique

Chapitre V : Analyse du corpus semi-spontané

Ce chapitre concerne les données qui forment le corpus « semi-spontané ». Elle présente l'expérience linguistique qui a été menée, l'analyse quantitative et les résultats que nous en avons tirés.

1. Méthodologie

1.1 Les tests réalisés

1.1.1 Introduction

DEF N DEF A fait l'objet d'une variation interne dans la langue et son apparition n'est pas fréquente. Celle-ci est très dépendante du contexte et des conditions d'énonciation. Afin d'accroître la taille de notre corpus, il a fallu provoquer l'usage de DEF N DEF A grâce à un test qui constitue les données semi-spontanées (de type (c), cf. section 3.4, p. 53). Un traitement systématique de la parole émise dans une situation contrôlée permet en effet une analyse optimale selon Stivers et Enfield (2010 : 2620). Il permet de collecter de manière systématique et comparable des données qui viennent compléter celles du discours spontané. L'inconvénient de ce dernier est en effet le caractère aléatoire de son contenu thématique et pragmatique, ce qui rend aléatoire également l'apparition d'une telle structure marquée, et donc nécessaire la collecte puis l'analyse d'un corpus immense. Au contraire, un test linguistique permet au linguiste de contrôler le contenu et les *stimuli*, le cadre de l'énonciation : « l'un des principes fondamentaux de l'expérience psycholinguistique est de décorréler les variables mises en jeu pour pouvoir estimer la valeur explicative de chaque variable sur la variable dépendante » (Thuilier 2012 : 94). On peut formuler des prédictions quant au type, à la fréquence, aux causes des structures susceptibles d'apparaître : les données viennent ensuite valider ou non son hypothèse préalable. Un tel corpus, créé dans des conditions contrôlées, permet une analyse plus fine, et également un plus grand nombre de données – même si le contexte d'apparition des occurrences est toujours le même, celui choisi pour la tâche.

Ne perdons pas de vue, en effet, que « la langue est un système abstrait dont le corpus est un produit : on ne peut pas réduire la langue au corpus » (Fagard 2015 : 11), et que la structure qui nous intéresse est optionnelle et non pas systématique. La probabilité qu'elle soit employée dans une heure d'enregistrement de corpus spontané, même d'un entretien dirigé ou semi-dirigé, est donc peu élevée. La Figure 36 illustre le rapport entre la langue et le corpus : la structure polydéfinie DEF N DEF A se trouve en romani dans le grand cercle (en pointillé) du « possible en langue »… mais seulement de manière sporadique et épisodique dans le premier cercle médian (en tireté) de l'« écrit / énoncé /

attesté ». Notre challenge consiste à faire apparaître la structure dans le deuxième cercle médian (en trait plein). Le fait de multiplier et de croiser les corpus (ainsi que les méthodologies : triangulation) permet de multiplier les occurrences présentes dans ce cercle, partant de mieux comprendre le fonctionnement de la structure qui nous intéresse. Enfin, le fait de « construire » les données, c'est-à-dire d'inciter les locuteurs à produire une parole semi-spontanée ou non-spontanée permet d'aborder plus facilement les phénomènes syntaxiques rares (Thuilier 2012 : 45).

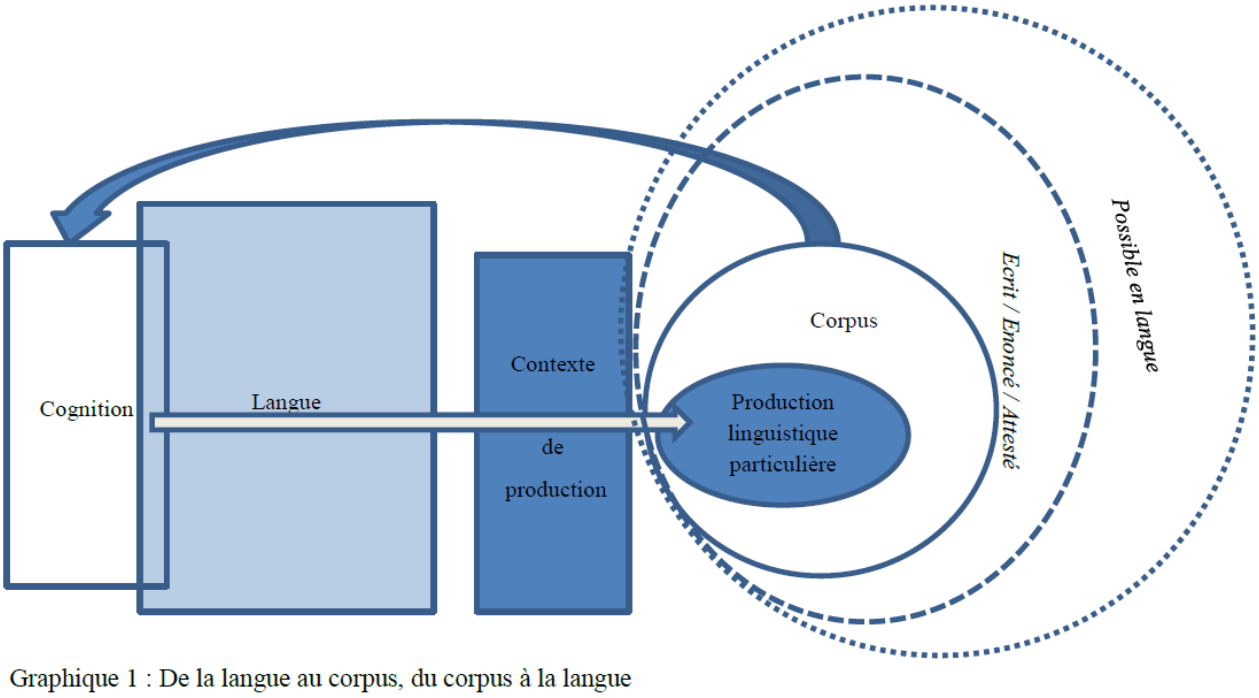


Figure 36 : De la langue au corpus, du corpus à la langue (Fagard 2015 : 11)

Il fallut donc placer les locuteurs dans un contexte jugé favorable à l'apparition de DEF N DEF A : le contexte retenu est celui du **focus à effet de sélection, avec un contraste *in presentia***. Pour déterminer ce contexte, nous nous sommes appuyée sur l'analyse que font entre autre Kolliakou (2004 : 268-271) et Lekakou et Szendrői (2012 : 125-129) du rôle de l'interprétation restrictive de l'adjectif épithète dans les structures polydéfinies (cf. section 2.4.2, p. 179). Nous partons en effet du postulat qu'il existe une corrélation entre la forme et la fonction. Une notion fonctionnelle est une raison nécessaire ou suffisante pour l'emploi d'une forme, ce que Skopeteas (2012 : 217) formalise de la manière suivante : « (1) *A property* E_i *of linguistic expressions occurs if a contextual condition* C_j *holds true* ».

Plutôt que d'inventer une situation verbale présentant un tel contexte, et dans l'optique de produire des données similaires entre elles donc comparables, nous nous sommes appuyée sur le *Questionnaire on Information Structure* (abrégé QUIS). Il s'agit d'un large dispositif expérimental visant à aider à la création d'une large base de données interlinguistique pour l'étude de la structure de l'information (Skopeteas et Féry 2007 : 327-328). Les expériences sont pensées pour révéler un aperçu des universaux et de la typologie des langues selon la façon dont elles encodent la structure de l'information :

- trois cent quatre-vingts phrases à faire traduire par les locuteurs de la langue d'étude
- un questionnaire grammatical « classique » sur les propriétés informationnelles de la langue d'étude,
- vingt-neuf tâches à réaliser en contexte écologique et discours spontané ou semi-spontané, pour faire ressortir telle ou telle propriété de la langue d'étude, le plus souvent au moyen de matériaux et de stimuli non-verbaux (images, vidéos muettes, objets, jeux).

Nous avons sélectionné une tâche au sein de ce dernier module et mené une micro-étude pilote en juillet 2014 avec Evangelia Adamou[1], locutrice native du grec moderne. Le contexte s'étant révélé convaincant, le test a été validé puis réalisé sur le terrain en juillet-août 2014 lors de ma deuxième mission en Albanie.

1.1.2 Test 1

1.1.2.1 Principe

Le test s'inspire de la tradition socio- et psycho-linguistique. Il s'agit d'un jeu d'instruction et de localisation statique, où un participant A donne des ordres à un participant B. Il s'inspire de la tâche n°8 du QUIS. Son principe est le suivant (Skopeteas et al. 2006 : 93) :

> « *Informant A sees a static powerpoint presentation in which animal cards are arranged in space. A second informant is given the cards shown in the presentation, but he does not see the presentation. The task of informant A is to give an accurate description of the arrangement in the presentation, so that informant B can reconstruct the same arrangement with his cards on a table. After each description informant A may check what informant B has done and may correct him if necessary.* » (Skopeteas et al. 2006 : 93)

L'intérêt de ce test est également « *the absence of priming as to which module(s) of grammar (prosody, word order, morphological markers) are to be used in a particular situation* » (Skopeteas et al. 2006 : 7). Elle fait naître une interaction variée entre les locuteurs : lexique spatial, verbes d'ordre, de défense, correction, questions d'alternative, de valeur de vérité, de précision, etc. Ce qui nous intéresse ici sont les différentes stratégies (prosodiques, morphologiques, syntaxiques, lexicales) adoptées par les locuteurs pour sélectionner, décrire et contraster les objets présents sur la table.

Le but était d'amener le participant A à nommer et localiser des objets sans pouvoir utiliser de déictiques, éventuellement d'amener le participant B à poser des questions ou demander des précisions. Les objets étaient disposés dans l'espace pour amener les participants à contraster des objets de nature identique (pomme de terre, stylo...) mais de propriétés différentes (taille, forme, aspect) ou bien à contraster des objets différents de propriétés identiques. Je n'ai jamais expliqué aux participants le but réel de l'expérience (leur faire produire DEF N DEF A), afin de ne pas influer sur leur

[1] Directrice de recherche au Lacito (UMR 7107 du CNRS), Villejuif.

choix des structures nominales[1]. Je parlais d'une étude comparative des différentes variétés albanaises du romani sur le plan de la prononciation, du lexique de la vie quotidienne (termes désignant des aliments) et des termes de localisation (« gauche », « droite », « aligné »).

Il s'agissait de réduire au minimum le biais linguistique en évitant ou réduisant l'utilisation d'une langue étrangère de contact (albanais, grec, italien) avant ou pendant la réalisation de la tâche ainsi que ma propre intervention. Le test est ainsi monolingue en romani.

1.1.2.2 Matériel

Une expérience a été conçue par Evangelia Adamou pour éliciter cette construction. Elle a adapté la tâche du QUIS (cf. section 1.1.1, p. 287) en utilisant des *realia*, des objets de la vie quotidienne (savon, boîte d'allumettes, bougie, chaise…) plutôt que des cartes et qu'une présentation Powerpoint. Elle suppose en effet que la présence immédiate d'objets en trois dimensions permet aux locuteurs une plus grande spontanéité que la visualisation d'images en deux dimensions, *a fortiori* sur un écran d'ordinateur. Elle vise à réduire la distance culturelle, notamment avec les plus âgés puisque les objets sont tirés de la réalité des locuteurs et immédiatement (re)connaissables. Elle vise aussi à réduire la difficulté de la tâche demandée et à faciliter la participation de l'ensemble des locuteurs, sans exclure ceux que l'utilisation de l'ordinateur pourrait incommoder. Enfin, l'utilisation d'objets nécessite moins de matériel et du matériel moins fragile. La réalisation du test est commode et rapide en toutes circonstances.

Chacun des participants disposait des quinze mêmes objets devant lui sur une table (Tableau 57). On peut regretter ici de n'avoir pas échappé à un certain biais de genre : la tâche ayant été préparée sur place avec l'aide de ma collaboratrice F16AČ, nous avons sélectionné un certain nombre d'objets relevant culturellement du genre féminin (baumes à lèvres, vernis à ongles, boucles d'oreille, bague, perle en plastique… pinces à linge). Il était nécessaire de trouver des objets suffisamment robustes et tous disponibles en double exemplaire. Les participants hommes ont souvent hésité pour nommer ces objets spécifiques, ce qui pourrait être dû au fait qu'ils manipulent peu, donc nomment peu ces objets dans leur vie quotidienne. Cela n'était pas nécessairement pertinent quant à l'utilisation de DEF N DEF A mais a pu contribuer à un sentiment d'insécurité linguistique lors du test.

Tableau 57 : Liste et caractéristiques des objets utilisés lors du test 1

Objet	Taille	Couleur	Aspect
pomme-de-terre	grosse		ovale partiellement pourrie
pomme-de-terre	moyenne		tordue
pomme-de-terre	petite		ronde

[1] Un tel détournement de l'attention est parfois nécessaire dans les expériences linguistiques ou psycho-linguistiques. Il peut faire partie intégrante du protocole, ainsi celui de l'expérience de la prononciation régionale et l'expérience d'instruction et de localisation avec les Lego chez Kehrein (2002).

Objet	Taille	Couleur	Aspect
vernis à ongles		blanc	
vernis à ongles		rose	
pince à linge		rose	
pince à linge		verte	
crayon		jaune	
crayon		noir	
crayon		vert	
crayon		rouge	
crayon	petit	orange	
stylo à bille		noir	
baume à lèvres		orange	
boucle d'oreille		verte et blanche	

1.1.2.3 Procédure

L'expérience sous sa forme définitive a consisté à réunir deux participants assis face à face à une table divisée en deux par un rideau tendu entre eux, de sorte que chacun voyait le visage mais pas la partie de table de l'autre. Chacun disposait des mêmes objets sur la table, prédisposés dans l'espace chez A (Figure 37) et laissés dans un ordre aléatoire chez B.

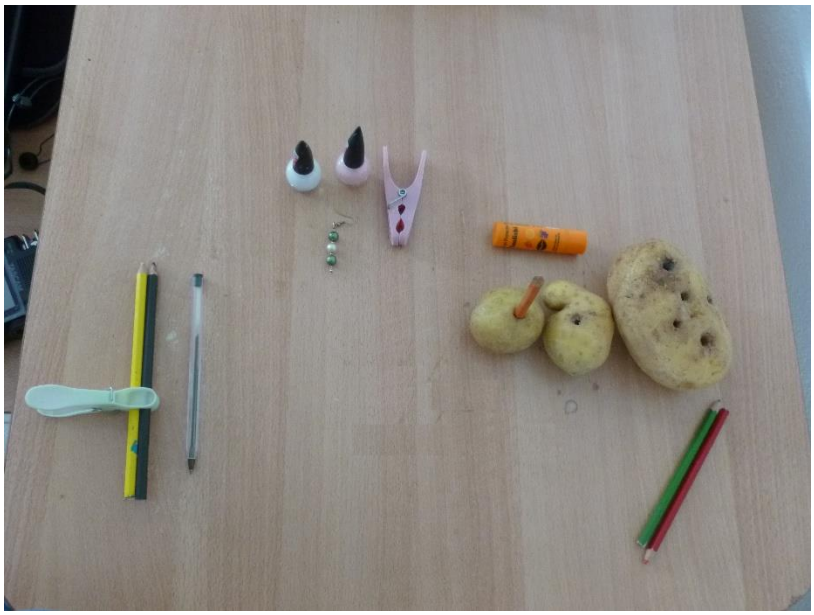


Figure 37 : Arrangement préalable des objets sur la table de A (test 1)

Le participant A devait décrire ce qu'il avait sous les yeux afin que le participant B puisse réorganiser les objets dans son espace comme ils se trouvent chez A (Figure 38). Une fois chaque objet mentionné par A et « placé » par B, le rideau était levé et je demandais à A de comparer les deux parties de la table, alors censées constituer des images miroir. Le participant A était censé détecter et commenter les différences et amener le participant B à ré-arranger les objets, afin d'établir l'image miroir.

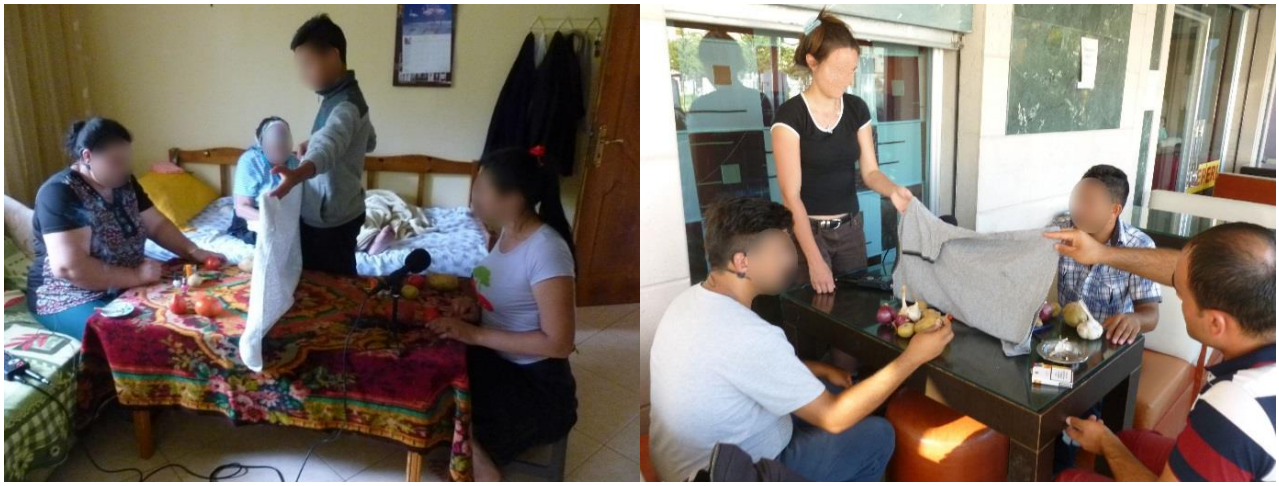


Figure 38 : Exécution du test 1

1.1.2.4 Enregistrement

Le test 1 représente seize enregistrements, soit 1h 1min 32sec (cf. Tableau 121). Un test durait en moyenne 3min 51sec. Les données ont été enregistrées en format WAV avec un enregistreur Tascam DR-100 alimenté à la fois par une batterie (rechargeable) et par des piles autonomes (rechargeables), très utiles en cas de séances de travail répétées ou prolongées. Le participant A portait un microphone serre-tête Shure Beta 54 et le participant B disposait d'un microphone statique Bey Dynamic M88 TG posé sur la table pour enregistrer ses éventuelles questions... et venir suppléer à un éventuel problème du microphone porté par A. En fin de séance, les fichiers étaient transférés de la carte de l'enregistreur à un ordinateur portable et copiés dans un disque dur externe. Je prenais en outre, après le test, des notes manuscrites dans un carnet : bien souvent les locuteurs avaient produit DEF N DEF A hors enregistrement, avant ou après le test, afin de commenter celui-ci.

1.1.2.5 Participants

Le test a été réalisé par 14 locuteurs natifs stratifiés par genre, âge, niveau d'éducation et variété (Les numéros notant les modalités des variables sociales sont expliqués dans le Tableau 123). Le test a été mené avec le souci de respecter la parité de genre (Tableau 58) et d'âge (

Tableau 59) :

Tableau 58 : Nombre de locuteurs ayant participé au test 1 en fonction du GENRE

	Femmes	**Hommes**	**- 15 ans**	**15 – 39 ans**	**40 – 59 ans**	**+ 60 ans**
	8	6	3	7	3	1
Total	14		14			

Tableau 59 : Nombre de locuteurs ayant participé au test 1 en fonction de la CLASSE D'AGE

	- 15 ans		15 – 39 ans		40 – 59 ans		+ 60 ans	
	Femmes	**Hommes**	**Femmes**	**Hommes**	**Femmes**	**Hommes**	**Femmes**	**Hommes**
	2	1	5	2	1	2	0	1
Total	3		7		3		1	

Il a été très difficile de réunir un échantillon à la fois sociologiquement représentatif de la communauté et quantitativement suffisant. En effet, il n'existe pas à notre connaissance d'étude sociologique sur la démographie de la population Rom albanaise (moyenne d'âge, répartition des classes d'âge, pyramide des âges en fonction du genre…), *a fortiori* une étude qui prenne en compte le niveau de formation, la profession, les groupes dialectaux. Nous ne pouvons donc que supposer qu'il existe autant de femmes et d'hommes dans l'ensemble de la population Rom. Nous ne pouvons que nous fier aux entretiens ethnographiques que nous avons menés, selon lesquels la population čergar I est moins nombreuse que celles mečkar et arli. Mais quelle est la proportion exacte des Mečkars par rapport aux Arlis ? Un tel échantillon « représentatif » l'est donc surtout des projections du chercheur sur la communauté.

En outre, il n'est pas facile, sur place, de trouver des personnes disponibles et acceptant de se prêter au jeu, surtout dans la mesure où la participation au test n'était pas rémunérée. Attirer les participants par la perspective d'une rémunération permet en effet au chercheur de sélectionner les participants en fonction de leur profil et donc d'atteindre la représentativité souhaitée, tandis que l'absence de rémunération oblige clairement le chercheur à se contenter des personnes volontaires et disponibles. Enfin, il est très difficile de réaliser une expérience représentative sur un petit corpus collecté dans un temps limité pour la réalisation du présent travail.

1.1.2.6 Difficultés rencontrées

J'ai rencontré des difficultés pour nommer le test. Dans les premiers temps, je l'ai appelé *test*, puis *ćhelibe*, jeu, qui signifie également « danse ». Pour éviter cela, j'ai utilisé ensuite le terme *loja*, « jeu » en albanais, qu'utilisaient déjà certains locuteurs. J'espérais que cette stratégie m'aiderait à convaincre un certain nombre de personnes réticentes à participer, en diminuant le stress de l'enregistrement audio, mais aussi l'impression de passer un « examen » (certaines personnes prétendaient ne pas « savoir la langue »). Mais le terme de « jeu », qui fait référence au monde des enfants, éloignait d'autres locuteurs, qui ne voulaient pas ou n'avaient pas le temps de s'abaisser à « jouer » – cette dénomination privait le test du respect conféré à une expérience scientifique. Le test lui-même souffrait d'un déficit de reconnaissance car il n'était pas perçu comme du « travail », puisque les personnes n'en comprenaient pas l'utilité – au contraire d'activités socialement reconnues et valorisées, telles que les entretiens ethnographiques qui s'apparentent à du travail journalistique de

transmission d'information. Bref, il m'était bien plus facile d'obtenir les métadonnées que les données elles-mêmes.

En outre, le test ne fonctionnait pas toujours comme je m'y attendais, et ce pour plusieurs raisons. Tout d'abord parce que les personnes jouant le rôle de A avaient parfois tendance à vouloir saisir et à déplacer les objets sur leur table en même temps que B, et donc à changer elles aussi la disposition des objets. Ensuite parce qu'il m'était presque impossible, une fois le rideau tombé, d'empêcher A d'utiliser des déictiques pour décrire les différences entre les deux parties de la table. La structure dont je voulais provoquer l'usage n'apparaissait donc quasiment jamais. Il était également difficile d'empêcher A de saisir lui-même les objets du participant B pour les déplacer, c'est-à-dire d'utiliser les mains plutôt que les mots. Ainsi, cette phase suivant le lever de rideau ne m'a pas semblé pertinente pour l'objectif qui était le mien.

Enfin, la phase finale, lorsque le « rideau » se levait, provoquait souvent incompréhension et désintérêt chez les participants. En effet, l'important était pour le participant A d'avoir décrit et manipulé les objets, d'avoir su les nommer, quoi que fît le participant B de son côté. Il arrivait souvent qu'il ne veuille pas ou ne parvienne pas à voir les différences entre les deux parties de la table, par désir de bien faire ou par peur d'offenser B. En effet, les personnes A et B éprouvaient parfois de la méfiance vis-à-vis de cette phase comme après un examen, qui place B dans la position d'un élève qui aurait éventuellement « mal fait » – alors que les différences entre les deux parties de la table pouvaient provenir aussi bien d'une description imprécise ou « erronée » de la part de A que d'un placement imprécis ou « erroné » de la part de B. S'ensuivait parfois un débat sur « à qui la faute », où je pouvais être mise en position de juge ou d'arbitre. Mon rôle n'étant pas de mettre mal à l'aise les locuteurs et la tâche n'étant en aucune façon un examen, j'ai préféré ne pas poursuivre de cette manière.

1.1.3 Test 2

1.1.3.1 Procédure

Dans une deuxième version du test, j'ai cessé de préétablir l'arrangement des objets du participant A. Je lui demandais simplement de les saisir et de les manipuler à sa guise, tout en décrivant ses gestes afin que le participant B puisse l'imiter de son côté. Cela a rendu les données moins comparables de façon systématique, mais a permis une augmentation du nombre d'occurrences produites, car les locuteurs se sentaient plus libres pour s'exprimer. Certains manipulaient plusieurs fois un même objet, ce qui augmentait le temps d'enregistrement. J'ai en outre supprimé la phrase finale de comparaison des deux parties de la table. Se sentant moins « surveillés », ils ont souvent laissé libre cours à leur créativité et réalisé des dispositions originales (Figure 39).

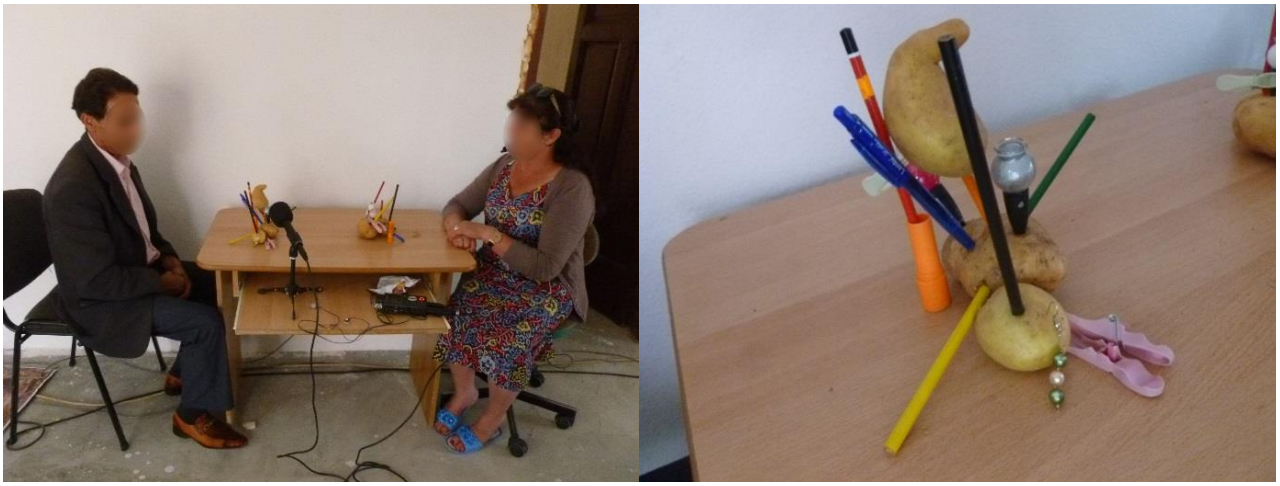


Figure 39 : Arrangement libre des objets sur la table de A (test 2)

1.1.3.2 Matériel

De nouveaux objets sont venus s'ajouter à ceux de la première version (Tableau 60).

Tableau 60 : Liste et caratéristiques des objets utilisés lors du test 2

Objet	Taille	Couleur	Aspect
pomme-de-terre	grosse		ovale partiellement pourrie
pomme-de-terre	moyenne		tordue
pomme-de-terre	petite		ronde
vernis à ongles		blanc	
vernis à ongles		rose	
pince à linge		rose	
pince à linge		verte	
crayon		jaune	
crayon		noir	
crayon		vert	
crayon		rouge	
crayon	petit	orange	
stylo à bille		noir	
baume à lèvres		orange	
boucle d'oreille		verte et blanche	
tomate	grosse		ronde
tomate	moyenne		pointue
tomate	petite		ronde
stylo à bille		bleu	
tête d'ail	grosse		
tête d'ail	petite		

Objet	Taille	Couleur	Aspect
oignon	gros		
oignon	petit		
bague	petite		
bague	grosse		
perle / bille		rouge	
prune	grosse		
prune	petite		
canette de soda			

1.1.3.3 Enregistrement

Le test 2 représente cinquante-six enregistrements (dont deux sont morcelés en plusieurs fichiers audio, suite à un problème technique), soit 3h 48min 18 sec (cf. Tableau 122). Un test durait en moyenne 4min 05sec.

1.1.3.4 Participants

Le test a été réalisé par trente-et-un locuteurs natifs stratifiés par genre, âge, niveau d'éducation et variété ainsi que par trois « témoins » (les numéros notant les modalités des variables sociales sont expliqués dans le Tableau 123). Ces derniers sont originaires d'autres pays de la péninsule balkanique et parlent d'autres variétés de romani que celles parlées en Albanie. Leur participation permettront d'en tirer des conclusions sur le plan dialectologique.

Tous les participants du test 1 (quatorze personnes) ont collaboré également au test 2 à l'exception de trois personnes, F9A, F11A et M60A. Onze personnes sur trente-sept ont donc passé les deux tests, trois n'ont fait que le test 1 et vingt-trois que le test 2. Le test a été mené avec le souci de respecter la parité de genre (Tableau 61) et d'âge (Tableau 62) :

Tableau 61 : Nombre de locuteurs ayant participé au test 2 en fonction du GENRE

	Femmes	Hommes	Témoins femmes	Témoins hommes
	16	15	1	2
Total	31		3	

Tableau 62 : Nombre de locuteurs ayant participé au test 2 en fonction de la CLASSE D'AGE

	- 15 ans		15 – 39 ans		40 – 59 ans		+ 60 ans		Témoins 15 – 39 ans		Témoins 40 – 59 ans	
	F	H	F	H	F	H	F	H	F	H	F	H
	0	2	9	7	3	5	4	1	1	1	0	1
Total	2		16		8		5		2		1	

Les tests ont rarement été réalisés de manière isolée. Sur trente-neuf « couples » de locuteurs des tests 1 et 2, onze ont été enregistrés en ma seule présence, mais vingt-huit en présence d'un ou de plusieurs tiers. C'est qu'ils avaient lieu le plus souvent au domicile de mes hôtes (chambre, salon, couloir, cour), au domicile d'autres participants ou au café. Presque toujours, d'autres membres de la famille ou des amis s'y trouvaient, écoutaient et commentaient ce qui se faisait – ne manquant pas d'accentuer le côté « mise en scène » de la tâche, donc d'amuser ou de stresser les participants. D'autres chercheurs ont été confrontés à ce phénomène, ainsi Adamou (2015 : 2-3), dont les enregistrements ont été réalisés dans la maison, la cour ou sur le lieu de travail des participants, et ce en présence de spectateurs, amis ou membres de la famille. Dans la communauté, la maison ou le lieu de travail sont en effet des lieux de passage habituel et ouvert, pour venir saluer, prendre des nouvelles, bavarder un moment. J'ai pu observer que la présence d'autres membres de la communauté pouvait être neutre (pour certains participants), avoir un effet positif (pour certains participants qui en profitaient pour exhiber leurs connaissances linguistiques ou prendre le public à partie, comme au spectacle), ou un effet négatif et inhibiteur (notamment chez les enfants et les jeunes gens en présence de leurs parents, ou chez certaines femmes en présence d'hommes).

Notons enfin qu'environ un tiers des personnes (douze) parle grec parce qu'il vit/a vécu ou travaille/a travaillé en Grèce ou avec des Grecs, et deux tiers (vingt-cinq) ne parlent pas grec (mais ont/avaient éventuellement, dans la famille, des parents qui parlent/parlaient le grec).

1.1.3.5 Difficultés rencontrées

J'ai rencontré beaucoup moins de difficultés avec cette version du test qu'avec la précédente. Dans les deux versions cependant, de nombreux participants ont, dans le rôle de A ou de B, recherché mon jugement et mon approbation : comment dit-on cela ? ai-je utilisé le bon mot ? ai-je bien placé l'objet ? / ai-je saisi le bon objet ? l'ai-je mis au bon endroit ? que veut dire A ? qu'est en train de faire A en ce moment, est-ce qu'il fait comme moi ? … alors même que j'avais précisé que je ne prendrais pas partie, qu'il n'y avait pas de « bonnes » et de « mauvaises » manières de s'exprimer. J'adoptais dans ce cas deux attitudes : soit la concentration sur un autre aspect de la tâche, censé requérir toute mon attention (contrôler l'enregistreur, la position des microphones, remettre en place les câbles, tenir le rideau, écrire dans le carnet), soit l'encouragement par des hochements de tête et des sourires, lorsque je sentais le participant inquiet ou inhibé. Je précisais aux locuteurs, le cas échéant, qu'ils pouvaient employer les termes qui leur étaient plus commodes, hérités ou empruntés (l'emprunt étant inévitable pour désigner un certains objets, tels que le vernis à ongles).

Je me suis efforcée de ne pas interrompre le discours du participant A, sauf lorsqu'un non-respect des consignes rendait inutile l'enregistrement même de la tâche : par exemple lorsque A brandissait un objet à hauteur de sa tête pour le montrer à B et donc éviter de le nommer ; lorsque A prenait plusieurs objets à la fois (par exemple tous les crayons ensemble, dans l'intention de les placer tous en même temps, ce qui empêchait de les contraster en fonction de leur couleur) ; enfin lorsque A ne

contrastait pas ou pas assez les objets et que je voyais le participant B, trop désemparé, peiner à reconstituer « à l'aveugle » les mouvements de A.

1.2 Dépouillement et codage

J'ai transcrit l'ensemble des syntagmes nominaux produits par les locuteurs jouant le rôle de A dans une même colonne d'un tableau sous format Calc. Les autres colonnes, trente-trois en tout, portent des informations sur la réalisation de la tâche et le profil des locuteurs (Figure 91) ainsi que des analyses linguistiques (Figure 92). Le fonctionnement par colonne permet par la suite de segmenter le corpus de toutes les manières possibles, en utilisant l'outil *filtre*. Le codage des données est détaillé dans le Tableau 124.

Les marqueurs de discours et interjections ont été transcrites dans la colonne SN prononcées mais ont été ignorées lors du codage et de l'analyse des parties du discours, restant « transparentes » pour la structure. Ce sont des cas de syntagmes polydéfinis avec rupture interne, présentant une insertion mais pas de changement intonatif : par exemple le nom propre Genti au vocatif dans le syntagme *i patàta Genti i tikni*, « la petite pomme de terre Genti ! ». Il rappelle l'exemple (195) du corpus spontané avec rupture interne. Nous l'avons codé DEF N (NP) DEF A, mais ensuite comptabilisé avec les structures DEF N DEF A.

2. Résultats : variables linguistiques

Les résultats seront tout d'abord présentés d'un point de vue linguistique (quels sont les facteurs *linguistiques* pertinents pour expliquer la variation, *comment* s'opère-t-elle ?), puis d'un point de vue sociologique (quels sont les facteurs *sociaux* pertinents pour expliquer la variation, *qui* la pratique ?). À cette fin, nous examinerons un certain nombre de variables dont le nom sera représenté en petites majuscules. Nous visons à établir si la variation est inter- ou intra-locuteurs, et en quoi elle réside : les variables qui s'avèrent déterminantes pourront alors être qualifiées de « facteurs ».

La tâche ayant été réalisée sous deux versions, baptisées test 1 et test 2, les résultats globaux seront d'abord présentés de manière séparée (section 2.1, p. 299), puis conjointe lors de l'examen des différentes variables pour des raisons de place (sections 2.2, 2.3 et 2.4, p. 302, p. 306, p. 329). Les données des témoins seront quant à elles présentées séparément, afin de pouvoir comparer les résultats des variétés albanaises du romani avec ceux de variétés balkaniques mais non albanaises.

Les syntagmes nominaux suivants seront recensés dans cette section :

- polydéfinis (DEF N DEF A)[1]
- monodéfinis avec A postposé (DEF N A)
- monodéfinis avec A antéposé (DEF A N)
- poly- ou monodéfinis complexes[2]

Nous n'avons pas tenu compte des syntagmes à proposition relative prédicative, qui sont en distribution complémentaire avec les syntagmes à adjectif épithète. Après avoir présenté les résultats des deux tests réalisés et des témoins, nous nous pencherons ici sur trois critères d'ordre linguistique : le type de déterminant qui peut être répété, l'ordre dans lequel se présentent les constituants nominaux et les autres répétitions qui peuvent se produire, et enfin le type d'adjectif qui est employé.

2.1 Distribution des syntagmes définis

Cette section présente les résultats des deux tests réalisés par les participants albanais et du test 2 réalisé par les témoins.

[1] Les trois catégories DEF N DEF A, DEF N A ET DEF A N comprennent ici des syntagmes référant non des objets présents sur la table, mais servant à les localiser, par exemple *ki ana i vaver*, de l'autre côté (litt. du côté de l'autre). Ces syntagmes ont été écartés, pour des raisons de place, des analyses présentées dans Tirard (2017).

[2] Cette catégorie a été écartée, pour des raisons de place, dans les analyses présentées dans Tirard (2017).

2.1.1 Test 1

La distribution des syntagmes nominaux définis incluant un adjectif épithète est représentée en Figure 40 (cf. Tableau 127). Nous avons exclu du calcul la stratégie concurrente que constitue l'adjectif attribut dans une proposition relative (« le crayon qui est rouge » au lieu de « le crayon rouge »).

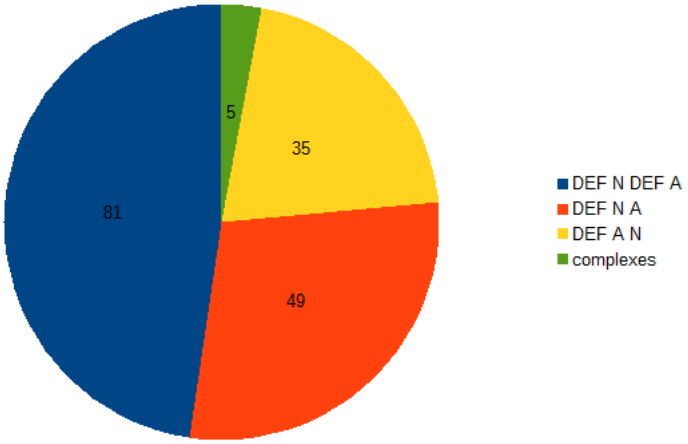


Figure 40 : Structure des SN définis produits par les participants du test 1

On trouve presque autant de monodéfinis (84 soit 49,41 %) que de polydéfinis (81 soit 47,65 %). Si l'on prend en compte les structures complexes qui sont pour la plupart des structures polydéfinies complexes, c'est presque l'égalité parfaite : 86 occurrences soit 50,59 %. La proportion de 47,65 % de structures DEF N DEF A est très importante, bien plus qu'en discours spontané. Cela montre que le choix de la tâche était pertinent et amenait en effet les locuteurs à utiliser plus de structures polydéfinies qu'à l'habitude. On note que parmi les syntagmes monodéfinis, on trouve une majorité d'adjectifs postposés (49, soit 28,82 %) et non d'adjectifs antéposés (35, soit 20,59 %).

2.1.2 Test 2

La distribution des syntagmes nominaux définis incluant un adjectif épithète est représentée en Figure 41 (cf. Tableau 127).

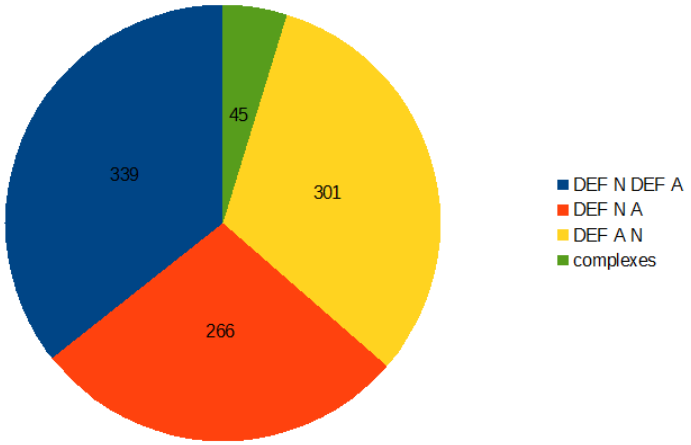


Figure 41 : Structure des SN définis produits par les participants du test 2

On trouve presque deux fois plus de monodéfinis (567 soit 59,62 %) que de polydéfinis (339 soit 35,65 %). Cette moindre proportion peut être symptomatique du caractère moins surveillé du discours dans cette version de la tâche, plus proche du discours spontané. Toutefois la proportion d'un tiers de polydéfinis reste beaucoup élevée qu'en discours spontané. Contrairement au résultat du test 1, on note que parmi les syntagmes monodéfinis, on trouve une majorité d'adjectifs antéposés (301, soit 31,65 %) et non d'adjectifs postposés (266, soit 27,97 %).

2.1.3 Tests 1 et 2

La distribution des syntagmes nominaux définis incluant un adjectif épithète est représentée en Figure 42 (cf. Tableau 127).

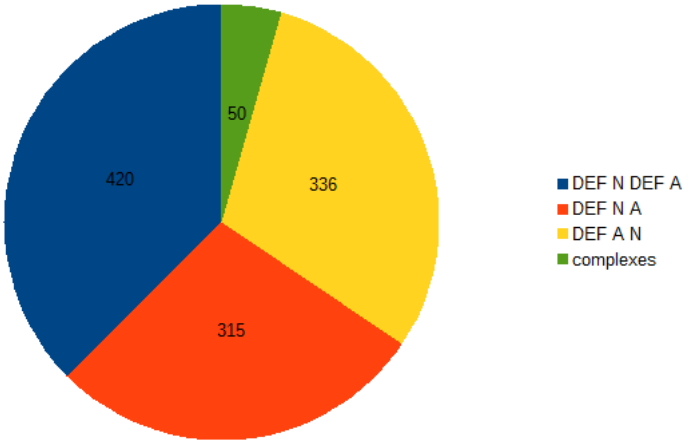


Figure 42 : Structure des SN définis produits par les participants des tests 1 et 2

On trouve une majorité de syntagmes monodéfinis (651 soit 58,07 %) par rapport aux polydéfinis (420 soit 37,47 %). On note que parmi les syntagmes monodéfinis, on trouve une courte majorité d'adjectifs antéposés (336, soit 29,97 %) et non d'adjectifs postposés (315, soit 28,10 %).

2.1.4 Témoins

La distribution des syntagmes nominaux définis incluant un adjectif épithète est représentée en Figure 43 (cf. Tableau 128).

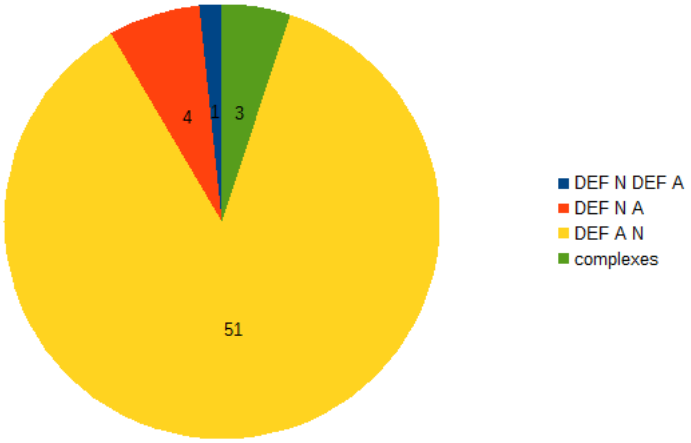


Figure 43 : Structure des SN définis produits par les témoins

On ne trouve qu'un seul polydéfini (soit 1,82 %) pour 51 monodéfinis (soit 92,72 %). L'écrasante majorité (51 occurrences, soit 86,44 %) des syntagmes est monodéfinie à adjectif antéposé.

2.1.5 Synthèse

La distribution des syntagmes nominaux définis incluant un adjectif épithète est globalement équilibrée en trois tiers (DEF N DEF A, DEF N A et DEN A N) dans l'ensemble des tests 1 et 2. La polydéfinitude est présente à 36,89 %, ce qui est un taux plus élevé qu'en discours spontané. On a constaté que le test sous sa première version amenait une plus grande production de syntagmes polydéfinis que sous sa seconde version, et que les témoins macédoniens et kosovar serbe pratiquent une syntaxe très différente de la moyenne des locuteurs albanais.

2.2 Variable 1 : DEFINITUDE

Il s'agit ici d'établir une typologie des déterminants qui peuvent être répétés, autrement dit de savoir si les locuteurs peuvent produire de la polyindéfinitude, de la polydémonstrativité, de la polypossessivité, ou si seule la poly*définitude*, la répétition de l'article défini, est possible.

2.2.1 Quel déterminant répéter ?

2.2.1.1 Tests 1 et 2

Dans le corpus du test 1 sont répétés dans un même syntagme les déterminants suivants (Tableau 63) :

Tableau 63 : Déterminants rédupliqués dans les tests 1 et 2

	DEF	IND	DEM	POSS	QUANT	NUM
Test 1	81	0	1	0	0	0
Test 2	339	0	0	0	0	0
Total	420	0	1	0	0	0

Le romani rappelle ici le grec tel qu'il était décrit dans la littérature : pas de réduplication de l'article indéfini (par exemple Androutsopoulou 2001 : 165-166). On note cependant la présence d'une structure « polydémonstrative » (sans pause ou changement d'intonation) :

(205) *M60A, Bilisht, 27 juillet 2014*

vətəm	*akale*	*lolipe*	*akale*	*tikn-e?*
seulement	DEM.PL	rouge_à_lèvres	DEM.PL	petit-PL

« seulement ces petits rouges à lèvres ? »

Ce syntagme constitue l'appendice d'une phrase peu articulée, dont le début est indistinct à nos oreilles comme à celles de notre collaboratrice principale (F16AČ, qui était la participante B). Le locuteur M60A demandait ici, en milieu de test, des précisions sur le fonctionnement de celui-ci. La situation d'énonciation est donc particulière.

2.2.1.2 Témoins

Dans le corpus des témoins, seul l'article défini est répété (Tableau 64) :

Tableau 64 : Déterminants rédupliqués par les témoins

	DEF	IND	DEM	POSS	QUANT	NUM
Occurrences	1	0	0	0	0	0

Ces données ne diffèrent pas de celles des locuteurs albanais, dans la mesure où aucun autre déterminant que l'article défini ne peut être répété.

2.2.2 Répéter l'article défini en présence d'un autre déterminant

2.2.2.1 Tests 1 et 2

La répétition de l'article défini s'accommode fort bien de la présence d'autres déterminants :

(206) *M45M, Vlorë, 13 août 2014*

odova/	*o*	*stilolapsi*	*o/*	*o*	*blu*
DEM.M.SG	ART.DEF.M.SG	stylo(M.SG)	ART.DEF.M.SG	ART.DEF.M.SG	bleu

« ce… ce stylo… bleu »

(207) *F60A, Korçë, 21 juillet 2014*

akova	*o*	*roza*	*o*	*manyćyri*
DEM.M.SG	ART.DEF.M.SG	rose	ART.DEF.M.SG	vernis_à_ongles(M.SG)

« ce vernis à ongles rose »

Ces syntagmes sont surdéterminés dans la mesure où ils présentent trois déterminants, un démonstratif et deux articles définis. Ils nous rappellent que le démonstratif peut exister de différentes manières : seul (208)a, antéposé à l'article défini et au nom (208)b, postposé à l'article défini et au nom (208)c, combiné à la polydéfinitude (208)d :

(208) a. *F60A, Llakatund, 13 août 2014*

odəja	*patata*
DEM.F.SG	pomme_de_terre(F.SG)
DEM	**N**

« cette pomme de terre »

b. *M45M, Llakatund, 13 août 2014*

akəja	*i*	*patate*
DEM.F.SG	ART.DEF.F.SG	pomme_de_terre(F.SG)
DEM	**DEF**	**N**

« cette pomme de terre »

c. *F47M, Llakatund, 13 août 2014*

i	*patata*	*odəja*
ART.DEF.F.SG	pomme_de_terre(F.SG)	DEM.F.SG
DEF	**N**	**DEM**

« cette pomme de terre »

d. *M45M, Llakatund, 13 août 2014*

odova/	*o*	*stilolapsi*	*o/*	*o*	*blu*
DEM.M.SG	ART.DEF.M.SG	stylo(M.SG)	ART.DEF.M.SG	ART.DEF.M.SG	bleu
DEM	**DEF**	**N**	**DEF**	**DEF**	**A**

« ce… ce stylo… bleu »

Cet état des lieux rappelle la description du romani d'Ajia Varvara par Igla (1996 : 165). Le démonstratif du romani d'Albanie (toutes variétés confondues) présente donc la même syntaxe que celui de cette variété. Il convient de noter que le grec autorise les trois constructions DEM DEF N (208)b, type 2 chez Igla), DEF N DEM (208)c type 3 chez Igla) et DEM DEF N DEF A (208)d non mentionné chez Igla ni ailleurs). Il s'agit d'une combinaison de (208)b et d'une structure polydéfinie classique. Chose remarquable de l'exemple (207), l'adjectif est antéposé au nom ([DEM] DEF A DEF N).

Notons en outre deux expressions polydéfinies avec POSS, dont nous n'avons pas tenu compte car elles ne s'appliquent pas directement aux objets du test :

(209) *F22M, Llakatund, 13 août 2014*

ko	*phiko*	*tir-o/*	*i*	*majti*
à.ART.DEF.M.SG	épaule(M.SG)	POSS.2SG-M.SG	ART.DEF	gauche

« à ton épaule/ gauche »

Ici l'adjectif est cependant emprunté à l'albanais : l'emprunt est lexical (le lexème *majti*), morpho-syntaxique (l'adjectif vient avec son article de connexion, *i* au lieu de *o* attendu ici) et syntaxique (l'adjectif est postposé au nom). Nous reviendrons sur cette question en section 2.4, p. 329.

2.2.2.2 Témoins

La répétition de l'article défini ne s'effectue pas en présence d'autres déterminants dans le corpus des témoins.

2.2.2.3 Synthèse

Dans les tests pratiqués, la répétition du déterminant se pratique en romani exclusivement avec l'article défini, si l'on excepte un cas de répétition du démonstratif. Il s'agit donc bien de polydéfinitude – le démonstratif étant conçu comme un déterminant défini complexe (Lyons 1999). On a pu remarquer que la polydéfinitude pouvait coexister avec la présence d'un autre déterminant (démonstratif) portant sur la même tête nominale. Nous n'avons pas d'exemple de poly*in*définitude dans ce corpus. On pourrait objecter que la nature du test amène à utiliser massivement l'article défini,

puisque les objets se trouvent littéralement sous la main des locuteurs. Cependant la situation contrastive favorable aurait dû amener la production de syntagmes polydéterminés avec d'autres déterminants si cela avait été autorisé par la langue. Parmi les 59 syntagmes indéfinis qui furent produits (Tableau 129 et Figure 93), on ne trouve notamment aucun syntagme poly*in*défini du type suivant (210) :

(210) **jek stilolapsi jek kal-o*

IND stylo(M.SG) IND noir-M.SG

« *un stylo noir »

2.3 Variable 2 : ORDRE DES CONSTITUANTS

Il s'agit ici d'établir quel est l'ordre des constituants nominaux dans le contexte communicatif de la tâche, quel que soit le déterminant qui introduit le syntagme. Nous verrons également si d'autres éléments que l'article défini peuvent être répétés au sein des syntagmes polydéfinis.

2.3.1 Syntagmes monodéfinis

2.3.1.1 Tests 1 et 2

L'adjectif peut se trouver dans deux positions différentes en syntagme monodéfini (Tableau 65).

Tableau 65 : Position de l'adjectif dans les syntagmes monodéfinis produits par les participants des tests 1 et 2

	DEF N A	DEF A N	Total
Test 1	49 (58,33 %)	35 (41,67 %)	84 (100 %)
Test 2	266 (46,91 %)	301 (53,09 %)	567 (100 %)
Total	315 (48,39 %)	336 (51,61 %)	651 (100 %)

La proportion d'antéposition et de postposition de l'adjectif est équivalente, même si l'on trouve une courte majorité d'adjectifs antéposés.

Lorsque le syntagme monodéfini présente plusieurs adjectifs, ils peuvent encadrer le nom (type DEF A N A) :

(211) *F66M, Voskopojë, 28 juillet 2014*

o	*žolto/*	*kalemi*	*žolto*
ART.DEF.M.SG	jaune	crayon	jaune
DEF	**A/**	**N**	**A**

« le jaune, crayon jaune »

(212) *F16AČ, Korçë, 19 juillet 2014*

o	*vaver*	*angarno*	*lol-o*
ART.DEF.M.SG	autre	crayon(M.SG)	rouge-M.SG
DEF	**A**	**N**	**A**

« l'autre crayon rouge »

(213) *F85A, Korçë, 1er août 2014*

i li	*i*	*vaver*	*sir*	*po*	*tikn-i*
FOC	ART.DEF.F.SG	autre	ail(F.SG)	COMP	petit-F.SG
	DEF	**A**	**N**		**A**

« et l'autre ail plus petit »

L'exemple (211) est un cas de reprise et d'hésitation. La locutrice a pu penser que l'adjectif seul ne suffisait pas à l'identification de l'objet, avant de se raviser, ne trouvant pas d'autre propriété à ajouter que celle déjà proposée, qui est alors répétée.

L'exemple (212) est prononcé après que la locutrice F16AČ vient de référer au crayon vert (DEF N A) à mettre près de la pomme de terre qui est pointue (DEF N SREL), puis le crayon qui est noir (DEF N SREL) à mettre près du crayon vert (DEF N A). Elle cite alors l'autre crayon rouge (DEF A N A) qu'il s'agira de mettre près du crayon noir (DEF N A). Elle prendra ensuite la pomme de terre noire (sic) (DEF N A) à mettre près du crayon rouge (DEF N A). On voit que dans l'exemple (212), les deux adjectifs sont contrastifs, puisque la couleur rouge permet de sélectionner un crayon parmi d'autres présents sur la table, de couleurs différentes mais déjà mentionnés en discours. L'adjectif « autre » n'a pas une portée longue puisqu'il porte seulement sur « crayon » et non sur « crayon rouge ». En fait, il n'y a qu'un seul crayon rouge sur la table, la sélection par {autre} s'effectue donc parmi l'ensemble {crayons} et non dans l'ensemble à l'intersection de {crayons} et de {entités rouges}. Or l'ensemble {crayons} n'existe pas sans propriété de couleur, il est donc impossible de réussir à sélectionner simplement un « autre crayon » (Figure 44) :

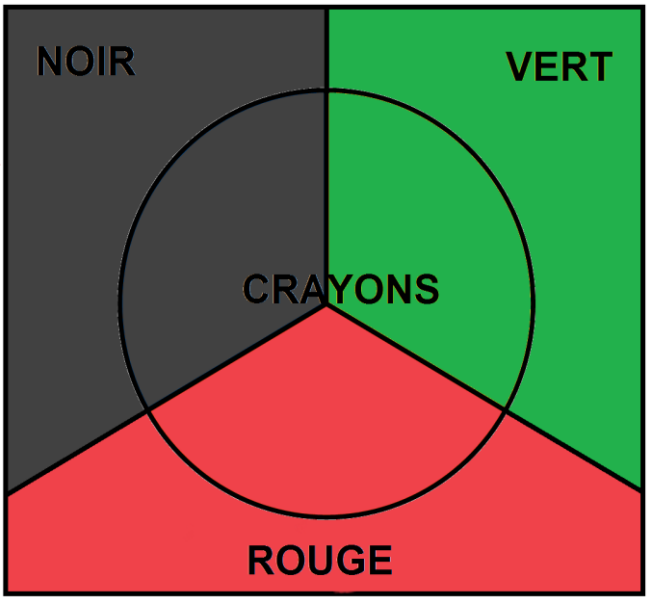


Figure 44 : Répartition de l'ensemble des crayons noirs, verts et rouges

Il n'existe pas de crayon qui ne soit pas caractérisé par une couleur, c'est pourquoi un adjectif de couleur suffit à identifier le référent. Même si son sens est intrinsèquement contrastif, on peut dire que l'adjectif *vaver*, autre, est l'adjectif le moins utile à l'identification, puisque, seul, il ne permettrait pas cette identification. Comme il s'agit du dernier crayon de la série qu'elle manipule, on peut même l'interpréter dans le sens de « non encore mentionné », voire dans le sens de « dernier ». Il s'agit d'une propriété informative plutôt que contrastive, ce qui est confirmé par sa position antéposée par rapport au nom.

L'exemple (213) est prononcé par la locutrice F85A alors qu'elle venait de mentionner une première fois l'ail, sans préciser plus avant ses propriétés (DEF N). Elle met ensuite le gros oignon (DEF A N) à côté de « l'ail (DEF N) là-bas au fond », sans préciser là non plus quel ail. Elle dit ensuite *ja i li i po tikni thava agore paš ti bari,* « oui et aussi le plus petit je le mets au fond à côté du grand ». Comme oignon et ail sont deux noms féminins, le genre des adjectifs *tikni* et *bari* ne permet pas de distinguer de quels objets il s'agit, d'où une demande de clarification de la part de la participante B. Elle reprend alors : *ja i li i vaver sir po tikni thava paš ti bari,* « oui et aussi l'autre petit ail je le mets à côté du grand ». On a ici une opposition binaire entre deux têtes d'ail, une grande et une petite. Cette opposition est portée à la fois par *vaver*, autre, puisque jusqu'à présent c'est la grande tête d'ail qui était censée être manipulée (même si la participante B n'avait pas, en fait, cette information), et par *po tikni,* plus petite, puisque cette propriété est directement opposée à *bari,* grande, au sein de la même phrase. Les deux adjectifs servent donc à l'identification, mais pas exactement de la même manière : l'adjectif antéposé (*vaver*) établit un contraste dans la situation (objet déjà manipulé ou non : *i sir* ~ *i vaver sir*, l'ail ~ l'autre ail) tandis que l'adjectif postposé, en outre mis au degré comparatif (*po tikni*), établit un contraste avec un autre adjectif au sein de la phrase, donc du discours (*i sir po tikni* ~ *i bari*, l'ail plus petit ~ le grand).

Nous n'avons pas trouvé d'exemple de modification multiple du nom, par plusieurs adjectifs qui se trouveraient du même côté du nom (DEF N A A ou bien DEF A A N), sauf les cas d'auto-correction précédés d'une pause. Dans ces cas, le deuxième adjectif vient annuler sémantiquement et pragmatiquement le premier :

(214) *M13AČ, Korçë, 7 août 2014*

ke	*o*	*manyćyri*	*parn-o/*	*roz*
FOC	ART.DEF.M.SG	vernis_à_ongles(M.SG)	blanc-M.SG	rose
	DEF	**N**	**A/**	**A**

« et le rouge à lèvres blanc, rose »

(215) *F65A, Korçë, 1er août 2014*

paš	*ti*	*xurd-i/*	*ti*	*bar-i*	*purum*
près	à.ART.DEF.F.SG	tout_petit-F.SG	à.ART.DEF.F.SG	grand-F.SG	oignon
	DEF	**A/**	**DEF**	**A**	**N**

« près du tout petit, du grand oignon »

Les syntagmes monodéfinis peuvent donc présenter un adjectif indifféremment pré- ou postposé, avec une courte majorité d'adjectifs antéposés.

2.3.1.2 Témoins

L'adjectif peut se trouver dans deux positions différentes en syntagme monodéfini (Tableau 66).

Tableau 66 : Position de l'adjectif dans les syntagmes monodéfinis produits par les témoins

	DEF N A	DEF A N	Total
F24K	2 (11,76 %)	15 (88,24 %)	17 (100 %)
M27Dž	2 (11,11 %)	16 (88,89 %)	18 (100 %)
M43X	0 (0 %)	20 (100 %)	20 (100 %)
Total	4 (7,27 %)	51 (92,73 %)	55 (100 %)

Nous avons trouvé un exemple de modification du nom par plusieurs adjectifs, tous antéposés :

(216) *F24K, Ohrid, 30 juillet 2014*

ko	*levo*	*naj*	*cikn-o*	*kompiri*
à.ART.DEF.M.SG	gauche	COMP	petit-M.SG	pomme_de_terre(M.SG)
DEF	**A**		**A**	**N**

« à la plus petite pomme de terre à gauche »

2.3.1.3 Répétition de l'affixe de cas

Que se passe-t-il lorsque le nom est fléchi, que ce soit à la strate I (cas oblique de certains référents animés, puisque les cas direct et vocatif ne sont pas marqués) ou à la strate II (cas génitif, datif, locatif ablatif, instrumental) ? Le corpus ne fournit pas d'exemple pour la strate I car les objets du test sont tous des référents inanimés. Il fournit en revanche quelques exemples pour la strate II, où l'adjectif est fléchi de manière canonique lorsqu'il est antéposé (c'est-à-dire que l'adjectif est fléchi à l'oblique de la strate I lorsque le nom est fléchi à la strate II, ici l'instrumental) :

(217) *M43X, Korçë, 20 juillet 2014*

e	*džungal-e*	*vast-e-sa*
ART.DEF.M.SG.OBL	dégoûtant-M.SG.OBL	main-M.SG.OBL-INSTR

« de la main gauche » [témoin]

(218) *M43X, Korçë, 20 juillet 2014*

e	*lač-e*	*vast-e-sa*
ART.DEF.M.SG.OBL	beau-M.SG.OBL	main-M.SG.OBL-INSTR

« de la main droite » [témoin]

Lorsqu'il est postposé, l'adjectif adopte le cas direct de la strate I :

(219) *M13AČ, Korçë, 7 août 2014*

paše	*e/*	*patat-a-ke/*	*bar-i*
près de	ART.DEF.M.SG.OBL	pomme_de_terre-F.SG.OBL-DAT	grand-F.SG

« à côté de/ la grande/ pomme de terre »

(220) *F22M, Llakatund, 13 août 2014*

kapse-ja-ha	*parn-i*
pince_à_linge-F.SG.OBL-INSTR	blanc-F.SG

« avec [la] pince à linge blanche »

(221) *M45M, Llakatund, 13 août 2014*

aćea	*laps-e-a*	*žolto*
DEM.F.SG.OBL	crayon-M.SG.OBL-INSTR	jaune

« avec ce crayon jaune »

(222) *M19M$_{(1)}$, Voskopojë, 28 juillet 2014*

e	*laps-e-sa*	*lol-o*
ART.DEF.M.SG.OBL	crayon-M.SG.OBL-INSTR	rouge-M.SG

« avec le crayon rouge »

L'adjectif est mis à une forme sous-spécifiée, non-marquée en cas. En (219), cette sous-spécification pourrait être due à la présence d'une pause entre le nom et l'adjectif. En (220), on pourrait penser que l'absence d'article défini perturbe l'économie générale du syntagme. En (221) et (222), les syntagmes sont introduits par un démonstratif et un article défini, et l'adjectif est bel et bien sous-spécifié.

2.3.2 Syntagmes polydéfinis

2.3.2.1 Un adjectif

L'adjectif peut se trouver dans deux positions différentes en syntagme polydéfini (Tableau 67).

Tableau 67 : Position de l'adjectif dans les syntagmes polydéfinis produits par les participants des tests 1 et 2

	DEF N DEF A	DEF A DEF N	Total
Test 1	81 (100 %)	0 (0 %)	81 (100 %)
Test 2	339 (99,41 %)	2 (0,59 %)	341 (100 %)
Total	420 (99,53 %)	2 (0,47 %)	422 (100 %)

Les deux syntagmes DEF A DEF N sont les suivants :

(207) *F60A, Korçë, 21 juillet 2014*

akova	*o*	*roza*	*o*	*manyćyri*
DEM.M.SG	ART.DEF.M.SG	rose	ART.DEF.M.SG	vernis_à_ongles(M.SG)

« ce vernis à ongles rose »

(223) *F66M, Voskopojë, 28 juillet 2014*

o	*zelino/*	*o*	*kalemi*
ART.DEF.M.SG	vert	ART.DEF.M.SG	crayon

« le crayon vert »

L'existence d'un tel ordre des constituants nominaux semble un phénomène marginal (0,48% des occurrences) et, puisque son existence n'est mentionnée nulle part dans la littérature, nous avons jugé, à ce stade, ne pas avoir de preuve suffisante de son existence. C'est pourquoi ces deux syntagmes n'ont pas été comptabilisés comme une catégorie à part dans le compte global de la distribution des syntagmes définis (Tableau 127 et calculs des variables sociales) mais l'ont été sous la catégorie *complexes*.[1]

[1] C'est également pourquoi ces deux occurrences ont été écartées des calculs et analyses présentés dans Tirard (2017).

Dans un troisième cas, celui de l'exemple (224), la pause rend le syntagme ambigu. Dans l'exemple (224), on hésite s'il faut considérer l'ensemble comme un syntagme polydéfini lâche à adjectif antéposé (DEF A/ DEF N) ou comme deux syntagmes différents juxtaposés :

(224) *M19M(2), Llakatund, 13 août 2014*

Le-ø	*li*	*i*	*ngrustik.*	*E*	*muk-ø*	*la/*
prendre-IMP.2SG	aussi	ART.DEF.F.SG	bague(F.SG)	eh	laisser-IMP.2SG	3.F.SG.OBL

ko/	*afər*	*manyćyri*	*o*	*lol-o,*	*le-ø*	*li*
à.ART.DEF.M.SG	près_de	vernis_à_ongles(M.SG)	ART.DEF.MSG	rouge-M.SG	prendre-IMP.2SG	aussi

o	***vaver/***	***i***	***ngrustik***	*čhiv-ø*	*la*	*nə*	*mest.*
ART.DEF.MSG	autre	ART.DEF.F.SG	bague(F.SG)	jeter-IMP.2SG	3.F.SG.OBL	dans	milieu

« Prends aussi la bague. Eh mets-la à/ près du rouge à lèvres rouge. Prends aussi l'autre/ la bague mets-la au milieu. »

Ici on peut douter que le syntagme soit polydéfini, car il y a non seulement pause, mais aussi changement de genre de l'article, comme si la locutrice voulait d'abord référer à un nom masculin (le dernier référent étant le vernis à ongles), puis se ravisait et référait finalement à un nom féminin (la bague). C'est pourquoi nous avons jugé qu'il s'agit d'une auto-correction, et qu'il y a ici deux syntagmes nominaux, pour deux référents différents.

2.3.2.2 Plusieurs adjectifs

Igla (1996 : 164-166) expliquait que la langue évitait la succession d'adjectifs antéposés ou postposés, et préférait dans ce cas encadrer le nom par des adjectifs. Cela se confirme dans notre corpus de syntagmes monodéfinis (225) :

(225) *F65A, F85A, Korçë, 1er août 2014*

o	*baro*	*kalemi/*	*o*	*saria*
ART.DEF.M.SG	grand-M.SG	crayon(M.SG)	ART.DEF.M.SG	jaune
DEF	**A**	**N/**	**DEF**	**A**

« le grand crayon/ jaune »

(226) *F65A, Korçë, 1er août 2014*

o	*kal-o*	*kalemi*	*o*	*bar-o*
ART.DEF.M.SG	noir-M.SG	crayon(M.SG)	ART.DEF.M.SG	grand-M.SG
DEF	**A**	**N**	**DEF**	**A**

« le grand crayon noir »

(227) *F44AČ, Korçë, 23 juillet 2014*

i	*vaver*	*domatka*	*i/*	*i*	*xurd-i*
ART.DEF.F.SG	autre	tomate(F.SG)	ART.DEF.F.SG	ART.DEF.F.SG	tout_petit-F.SG
DEF	**A**	**N**	**DEF**		**A**

« l'autre tomate/ toute petite »

(228) *F85A, Korçë, 1er août 2014*

i	*vaver*	*kurtulka*	*i/*	*i*	*po*	*tikin-i*
ART.DEF.F.SG	autre	pomme de terre(F.SG)	ART.DEF.F.SG	ART.DEF.F.SG	COMP	petit-F.SG
DEF	**A**	**N**	**DEF**			**A**

« l'autre pomme de terre/ plus petite »

Lorsque la locutrice F65A prononce (225), elle a mentionné auparavant le crayon rouge (DEF A N) à mettre entre deux pommes de terre. Elle saisit ensuite o saria, le jaune, (DEF A) puis veut préciser de quel objet il s'agit, estimant que la propriété de couleur ne suffit pas à l'identifier. Elle ajoute donc un nom et prononce o kalemi o/ (DEF A DEF/) mais change de propriété identificatoire puisqu'elle énonce o xurdo kalemi, le tout petit crayon, qu'il s'agit de mettre à côté de la toute petite pomme de terre (DEF A N). Comme il y a sur la table un (grand) crayon jaune, qui pourrait être confondu avec celui, tout petit, auquel elle réfère (qui est d'ailleurs orange, différence qu'elle n'effectue pas), elle choisit de l'identifier au moyen de la taille (puisqu'il est en revanche l'unique crayon tout petit sur la table). L'opération suivante consiste à prendre un autre objet, le grand crayon jaune. Pour elle, ce référent se distingue de l'autre non par sa couleur, puisque tous deux sont jaunes, mais par sa taille : le précédent était tout petit, celui-ci est grand. L'adjectif qui contraste directement est donc l'adjectif antéposé, de taille, contrairement à l'ordre des mots constaté dans l'exemple (212). Pour autant, on ne peut pas dire que l'adjectif postposé, de couleur ne contraste pas : juste après avoir posé le crayon en question à côté de la toute petite pomme de terre (DEF A N/ DEF A) elle saisit le crayon vert (DEF A N) qu'elle pose à côté du gros oignon (DEF A N), avant de saisir le crayon noir (def a n). Or les crayons vert et noir sont eux aussi de grande taille. On peut donc dire que l'adjectif postposé, dans « le grand crayon/ jaune », contraste avec d'autres référents présents dans la situation, mais non encore mentionnés. Il a une portée longue puisqu'il porte à la fois sur « grand » et sur « crayon ». En d'autres termes, la sélection par {jaune} s'effectue non parmi l'ensemble {crayons}, mais parmi l'ensemble déjà à l'intersection de {crayons} et de {entités grandes}.

On a vu que le crayon noir est mentionné une première fois (DEF A N) : la locutrice est alors interrompue par la participante B qui est en retard puisqu'elle n'a pas encore posé le crayon vert à côté *du gros oignon*. B dit pour elle-même *paš ti purum/ i bari*, à côté de l'oignon/ gros, ce qui fait comprendre à A qu'elle doit marquer une pause pour l'attendre. Lorsque A reprend le fil de son discours, elle fait de nouveau référence au crayon noir, en prononçant le syntagme (226). De son point de vue à elle, les deux adjectifs « noir » et « grand » contrastent puisqu'elle a déjà parlé de crayons d'autres couleurs (jaunes), dont un de grande taille… et elle a même déjà mentionné le crayon noir. Cependant, on l'a vu, de nombreux grands crayons contrastent entre eux par la couleur : c'est donc l'adjectif de couleur le plus utile à l'identification, et non l'adjectif de taille (d'autant que la

participante B n'était pas attentive au moment de la première mention du crayon noir). Mais dans le syntagme (226), l'adjectif de couleur n'est pas mis dans la position focale (postposée) mais dans la position informative moins contrastive (antéposée). Autrement dit, l'information attendue n'est pas (assez) saillante. Cela peut expliquer pourquoi la locutrice B interrompt la locutrice A juste après l'énonciation de (226), pour demander : *so nđyra?*, « quelle couleur ? ». A répond en référant à nouveau à l'objet : *o kalo kalemi*, le crayon noir (DEF A N), et répète même *kalo kalemi* (A N). On constate qu'elle continue à antéposer l'adjectif identificatoire, mais la répétition permet à B de comprendre l'information. Notons que la participante A, F65A, réticente à postposer l'adjectif en syntagme mono- ou polydéfini, a plus de 60 ans et parle la variété arli, tandis que la participante B, F44AČ, a entre 40 et 59 ans et parle la variété Arli & çergar. Nous verrons plus loin en quoi l'âge et la variété sont des facteurs pertinents pour expliquer l'organisation différente de l'information au sein du syntagme nominal. L'exemple (226) est l'un des rares syntagmes présentant plusieurs adjectifs qui ne respecte pas l'ordre canonique d'apparition des adjectifs épithètes (Stavrou 1999 : 202-203, Laenzlinger 2000 : 59-60) (cf. Figure 10 et Figure 12). L'adjectif de taille y est prononcé après l'adjectif de couleur – contrairement aux exemples (225) et (231).

Lorsque la locutrice F44AČ prononce l'exemple (227), elle a déjà mentionné le référent en question : elle a commencé l'enregistrement en saisissant cet objet (la grosse tomate, DEF N DEF A). Elle manipule ensuite divers objets (pommes de terre, vernis à ongles, crayons), dont la tomate pointue (DEF A), puis à nouveau la grosse tomate rouge (P.DEF N P.DEF A/ P.DEF A/ P.DEF A, exemple (250)). Elle saisit ensuite la boucle d'oreille qu'il s'agit de mettre à côté de la tomate (DEF N). Les deux ayant été manipulées, le participant B, M13AČ, lui demande si elle fait référence à la tomate pointue, ce à quoi A répond *na, i bari domata*, non, la grosse tomate (DEF A N). Elle enchaîne en disant *ka las i vaver domatka i/ i xurdi*, on va prendre l'autre tomate/ toute petite. Or il y a trois tomates sur la table : une grande, une moyenne pointue, et une (toute) petite (Tableau 60). *Vaver*, autre, n'est pas utile à l'identification puisque, pour contraster avec la grande, il peut renvoyer soit à la moyenne pointue, soit à la (toute) petite. Son statut est donc, comme dans l'exemple (212), plutôt informatif. L'adjectif utile à l'identification est celui postposé accompagné d'un article[1] : *xurdi*, toute petite.

L'exemple (228) est prononcé en début d'enregistrement. F20A déclare : *Le i kultulka i bari ka čhuves la maš/ maškare. Le i vaver kultulka i/ i po tikni ka čhuves la paš ti bari*, Prends la grosse pomme de terre (DEF N DEF A) tu vas la mettre au mi/ au milieu. Prends l'autre pomme de terre/ plus petite (DEF A N DEF/ DEF COMP A) tu vas la mettre à côté de la grande. Là encore, l'opposition sémantique binaire qu'offre *vaver*, autre, ne suffit pas à l'identification, d'où l'ajout d'un adjectif permettant l'identification, *tikni*, petite (au comparatif), opposé à *bari*, grande. Notons que le troisième adjectif servant à l'identification ne sera pas *xurdi*, tout petite, comme on pourrait s'y attendre, mais *banđi*, tordue. Cela montre que les locuteurs découpent la réalité des pommes de terre (Tableau 60) de manière différente : là où certains voient une grande (*bari*) ~ une petite (*tikni*) ~ une

[1] Le syntagme présente une pause située après le deuxième article défini, c'est-à-dire au milieu de la seconde sous-partie du syntagme global (non entre les deux sous-parties). C'est pourquoi il ne s'agit pas d'une *loose apposition*, apposition lâche, mais d'une *close apposition*, apposition étroite (*close polydefinite*), au sens de Lekakou et Szendrői (2007). L'adjectif postposé et l'article défini qui l'accompagne sont des modifieurs épithètes du nom *kurtulka*, pomme de terre, au sein du même syntagme nominal. La locutrice hésite sur l'adjectif exact à employer, donc sur la manière de qualifier la pomme de terre, et non sur l'emploi de la polydéfinitude.

toute petite (*xurdi*), d'autres voient une grande (*bari*) ~ une tordue (*banđi*) ~ une petite (*tikni*). L'adjectif *tikni*, petite, peut donc soit s'opposer trinairement à *bari* (grande) et à *xurdi* (toute petite) – sur une échelle de taille ; soit s'opposer binairement à *bari* (grande), si la pomme de terre moyenne est quant à elle définie par son aspect (*banđi*, tordue) – sur une échelle mêlant donc taille et aspect.

Il est toutefois possible, grâce à la polydéfinitude, de trouver une suite d'adjectifs postposés – ce qui contredit l'affirmation d'Igla (1996 : 166). On trouve dans le corpus polydéfini des exemples de modification multiple du nom DEF N A/ DEF A et DEF N DEF A A.

(229) *F22M, Llakatund, 13 août 2014*

i	*sir*	*vaver/*	*i*	*bar-i*
ART.DEF.F.SG	ail(F.SG)	autre	ART.DEF.F.SG	grand-F.SG
DEF	**N**	**A/**	**DEF**	**A**

« l'autre ail/ grand »

(230) *F22M, Llakatund, 13 août 2014*

o	*lapsi/*	*o*	*vaver/*	*žolto*
ART.DEF.M.SG	crayon(M.SG)	ART.DEF.M.SG	autre	jaune
DEF	**N/**	**DEF**	**A/**	**A**

« l'autre/ crayon/ jaune »

(231) *F20A, Korçë, 7 août 2014*

o	*lapsi*	*o*	*tikn-o*	*saria*
ART.DEF.M.SG	crayon(M.SG)	ART.DEF.M.SG	petit-M.SG	jaune
DEF	**N**	**DEF**	**A**	**A**

« le petit crayon jaune »

L'exemple (229) est prononcé par la locutrice F22M alors qu'elle a déjà mentionné six fois « l'ail » (DEF N) et deux fois « la petite tête d'ail » (DEF N DEF A). Elle fait alors référence à « l'autre tête d'ail » (DEF N A/ DEF A), qui contraste binairement puisqu'il n'y a que deux têtes d'ail sur la table: *vaver* suffit à l'identification. Elle prononce d'ailleurs une première fois un syntagme *i sir vaver*, l'autre ail (DEF N A), avant de prononcer le syntagme (229) *i sir vaver/ i bari*, l'autre ail/ grand (DEF N A/ DEF A). La répétition du syntagme ainsi que la pause, au sein du deuxième syntagme, entre *i sir vaver* et *i bari*, mais aussi la pause qui suit le syntagme (229) et l'intonation, laissent à penser que la locutrice cherche en réalité à gagner du temps pour trouver quoi faire de l'ail en question (elle saisit finalement une pince à linge pour l'accrocher à la queue de l'ail). Le premier adjectif est le plus utile à l'identification – non que le deuxième ne soit pas contrastif (les deux têtes d'ail s'opposent par la taille), mais il est pragmatiquement redondant. Comme dans l'exemple (213), *vaver*, autre, établit un contraste dans la situation (objet déjà manipulé ou non) tandis que l'adjectif de taille établit un contraste au sein du discours (*petit ~ grand*).

L'exemple (230) est prononcé juste après, alors que F22M a déjà mentionné deux fois le crayon jaune puis trois fois le crayon noir. Elle veut maintenant faire à nouveau référence au crayon jaune : dans son discours ils sont jusqu'à présent en opposition binaire, donc l'adjectif *vaver*, autre, est suffisant pour identifier le référent. Comme il y a cependant d'autres crayons sur la table (qu'elle ne mentionne que bien plus tard), elle prend la peine de préciser la couleur du référent, en deuxième position et sans article défini – ce qui laisse à penser que cette propriété est pragmatiquement secondaire. Cela peut expliquer pourquoi l'ordre des adjectifs ne respecte pas celui canonique présenté par épithète (Stavrou 1999 : 202-203, Laenzlinger 2000 : 59-60) (cf. Figure 10 et Figure 12). L'adjectif autre, orienté vers le locuteur, y est prononcé après l'adjectif de taille – contrairement aux exemples (227), (228) et (229).

L'exemple (231) est prononcé en début d'enregistrement par F20A et constitue la première mention d'un crayon. Il est intéressant de constater que ce n'est pas n'importe quel crayon auquel il est fait référence ici, mais à celui petit et orange (qualifié par elle de jaune). Il contraste avec les autres à la fois par sa taille et par sa couleur, d'où la double postposition de l'adjectif. C'est par sa taille qu'il contraste binairement avec le grand crayon jaune, d'où, peut-être la priorité donnée à *tikno*, petit, cité en premier et accompagné d'un article défini.

Dans les trois exemples qui précèdent, les adjectifs postposés ne sont pas tous accompagnés d'un article défini. Mais ce peut être le cas, ainsi en (232) – (236), structures polydéfinies à trois articles :

(232) *M18Č, Korçë, 24 juillet 2014*

i	*kapsa*	*i*	*vaver*	*i*	*lol-i*
ART.DEF.F.SG	pince_à_linge(F.SG)	ART.DEF.F.SG	autre	ART.DEF.F.SG	jaune-F.SG
DEF	**N**	**DEF**	**A**	**DEF**	**A**

« l'autre pince à linge rouge »

(233) *F16AČ, Korçë, 7 août 2014*

i	*kultulka*	*i*	*tikn-i*	*i*	*vaver*
ART.DEF.F.SG	pomme_de_terre(F.SG)	ART.DEF.F.SG	petit-F.SG	ART.DEF.F.SG	autre
DEF	**N**	**DEF**	**A**	**DEF**	**A**

« l'autre petite pomme de terre »

(234) *F16AČ, Korçë, 21 juillet 2014*

o	*manyćyr*	*o*	*vaver/*	*o*	*roza*
ART.DEF.M.SG	vernis_à_ongles(M.SG)	ART.DEF.M.SG	autre	ART.DEF.M.SG	rose
DEF	**N**	**DEF**	**A/**	**DEF**	**A**

« l'autre vernis à ongles/ rose »

(235) *F16AČ, Korçë, 24 juillet 2014*

o [sic]	*kapsa*	*i*	*vaver/*	*i*	*ješili*
ART.DEF.M.SG	pince_à_linge(F.SG)	ART.DEF.F.SG	autre	ART.DEF.F.SG	vert

DEF	**N**	**DEF**	**A/**	**DEF**	**A**

« l'autre pince à linge verte »

(236) *M19M(2), Llakatund, 13 août 2014*

I	*patata*	*i*	*tikn-i/*	*i*	*čhin-d-i*
ART.DEF.F.SG	pomme_de_terre(F.SG)	ART.DEF.F.SG	petit-F.SG	ART.DEF.F.SG	couper-PTCP-F.SG
DEF	**N**	**DEF**	**A/**	**DEF**	**A**

« la petite pomme de terre coupée »

Les exemples (232) et (233) ne présentent pas de pause, contrairement à (234), (235) et (236). *Vaver*, autre, est le premier des deux adjectifs dans les exemples (232), (234), (235) et le deuxième dans l'exemple (233). Comment peut-on l'expliquer ? C'est que les exemples (232), (234) et (235) font référence à des objets qui contrastent de manière binaire avec un seul autre (pince à linge et vernis à ongle) : l'adjectif *vaver*, qui n'a pas d'autre sens que le contraste, semble le plus approprié, voire cognitivement le plus économique, ce qui explique qu'il soit placé en première position. En effet, il n'y a que deux pinces à linge, donc l'autre est forcément la rouge, et la rouge est forcément l'autre : pas besoin de chercher à détailler une propriété intrinsèque du référent telle que la couleur, « autre » suffit, pas besoin non plus de chercher comment exprimer telle ou telle nuance de couleur qui n'a pas d'équivalent romani dans le stock lexical hérité. Cela explique que *vaver*, autre, soit antéposé en DEF A N dans seulement 14 occurrences, mais postposé en DEF N DEF A dans 19 occurrences, et en DEF N A dans 16 occurrences (sans compter 16 cas de complexes) : les 14 antépositions ne représentent que 28,57 % des cas. C'est pourquoi l'adjectif de couleur qui le suit est pragmatiquement redondant. On retrouve le même schéma qu'en (229), à ceci près que *vaver* est accompagné d'un article défini.

L'exemple (233) fait quant à lui référence à une pomme de terre, qui contraste de manière trinaire (cf. analyse de l'exemple (228) : *vaver* n'est pas indiqué pour identifier le référent à coup sûr – ce d'autant qu'à ce moment du discours, la locutrice F16AČ a déjà mentionné toutes les pommes de terre disponibles : trois fois la « grande » (DEF N/ DEF A, DEF N SREL, DEF N SREL), une fois la « tordue » (DEF N DEF A), une fois la « petite » (DEF N DEF A). Lorsqu'elle prononce *i kultulka i vaver*, l'autre pomme de terre, la participante B reste perplexe : il peut aussi bien s'agir de la petite que de la grande (la tordue vient juste d'être mentionnée). A ne finit pas sa phrase et se ravise, prononçant cette phrase énigmatique : *Ka ə les i kultulka i tikni i vaver ka čhuves la/ paš ti kultulka ti tikni/ ti vaver,* tu euh tu vas prendre l'autre petite pomme de terre et tu vas la mettre/ à côté de la petite pomme de terre, l'autre. La locutrice B n'a pas bronché ici mais a continué à exécuter la tâche de son mieux. Cette difficulté à distinguer une autre petite pomme de terre d'une autre petite pomme de terre vient d'une fusion de deux découpages de la réalité, où l'adjectif « petit » peut référer aussi bien à la pomme de terre n°2 qu'à la pomme de terre n°3. Lorsqu'il réfère à la pomme de terre n°3 cependant, c'est parce que l'adjectif « tordu » se charge de la n°2, comme elle fait elle-même au début.

On trouve également des structures polydéfinies à trois articles, mais dont l'un est albanais :

(237) *F47M, Llakatund, 13 août 2014*

i	*patata*	*i/*	*i*	*tikin-i*	*e*	*mesmja*
ART.DEF.F.SG	pomme_de_terre(F.SG)	ART.DEF.F.SG	ART.DEF.F.SG	petit-F.SG	ART.DEF.F.SG	moyenne
DEF	**N**	**DEF/**	**DEF**	**A**	**DEF**	**A**

« la pomme de terre/ petite moyenne »

(238) *F44AČ, Korçë, 23 juillet 2014*

o	*kalemi*	*o*	*xurd-o/*	*i*	*fundi*
ART.DEF.M.SG	crayon(M.SG)	ART.DEF.M.SG	tout_petit-M.SG	ART.DEF.M.SG	dernier
DEF	**N**	**DEF**	**A/**	**DEF**	**A**

« le tout petit dernier crayon »

La présence de la pause entre l'article et l'adjectif en (237) montre que la locutrice hésite sur la manière de qualifier l'objet (sur l'adjectif qu'elle s'apprête à prononcer), non sur l'emploi de la structure polydéfinie. Dans les exemples (237) et (238), le deuxième adjectif est emprunté à l'albanais : l'emprunt est lexical (le lexème *mesmja* ou *fundi*), morpho-syntaxique (l'adjectif vient avec son article de connexion, *e* au lieu de *i* attendu ici en romani, ou *i* au lieu de *o*) et syntaxique (l'adjectif est postposé au nom). Nous reviendrons sur cette question en section 2.4, p. 329.

Notons qu'en (237) les deux adjectifs qui modifient le nom ont un sens qui pourrait être considéré contradictoire : le référent est à la fois qualifié de petit et de moyen. On pourrait croire qu'il y a, pour elle, une quadripartition des adjectifs de taille (*baro/mesmja/tikno/xurdo*, grand/moyen/petit/tout-petit), mais il s'agit en réalité plutôt une tripartition, où *mesmja* et *tikno*, moyen et petit, sont ici équivalents, parce qu'opposés à *baro* et à *xurdo*, grand et tout petit. Notons que les trois pôles sont exprimés par les adjectifs hérités.

L'exemple (238) présente une pause après la première sous-partie adjectivale, et avant la deuxième. La locutrice hésite sur la propriété supplémentaire à mentionner pour mieux aider l'autre à identifier le référent. Les adjectifs de cet exemple ne respectent l'ordre canonique d'apparition des adjectifs épithètes (Stavrou 1999 : 202-203, Laenzlinger 2000 : 59-60) (cf. Figure 10 et Figure 12). Ceci peut être dû au fait que *fundi*, dernier, soit emprunté à l'albanais, langue qui postpose les adjectifs. Il ne peut pas se trouver en structure DEF A N. S'il est ressenti comme « étranger » au système grammatical, sa position sera éloignée du cœur du syntagme, le nom.

L'ordre que l'on relève dans les suites d'adjectifs respecte les typologies de (Stavrou 1999 : 202-203, Laenzlinger 2000 : 59-60) (cf. Figure 10 et Figure 12) et (Cinque 2010 : 38) (Figure 9) pour des langues à antéposition de l'adjectif (anglais, chinois) et des langues à postposition de l'adjectif (langues celtiques). (Stavrou 1999 : 215-216) observait que l'ordre canonique peut être violé en cas de postposition de l'adjectif en grec : ce n'est pas le cas dans notre corpus de romani. Voici l'ordre des adjectifs tel qu'on l'a relevé dans cette section :

Postposition : autre[1] > taille > couleur / aspect

Les adjectifs des syntagmes (229), (231), (232), (233), (234), (235), (236), (237) et (238) présentent une relation de coordination avec le nom et les adjectifs sont tous restrictifs. Dans le syntagme (230), il y a combinaison de subordination et de coordination : *vaver* est subordonné et *žolto* est coordonné : *lapsi vaver* n'est pas suffisant pour assurer l'identification, puisqu'il y a plusieurs autres crayons, d'où la nécessité de préciser sa couleur. Du point de vue structurel, il s'agit donc d'un [[autre crayon] jaune]. Le syntagme (229) présente au contraire deux adjectifs coordonnés, qui suffisent chacun à assurer l'identification, puisqu'il n'y a qu'un autre ail, à la fois autre et grand. Du point de vue structurel, il s'agit donc d'un [ail [autre, grand]] (cf. Nølke 2001, section 2.2.2.5, p. 79).

2.3.2.3 Répétition de l'affixe casuel

Que se passe-t-il lorsque le nom est fléchi, que ce soit à la strate I (cas oblique de certains référents animés, puisque les cas direct et vocatif ne sont pas marqués) ou à la strate II (cas génitif, datif, locatif ablatif, instrumental) ? Le corpus ne fournit pas d'exemple pour la strate I car les objets du test sont tous des référents inanimés. Il fournit en revanche deux exemples pour la strate II, où la marque du cas nominal est répétée sur l'adjectif :

(239) *F22M, Llakatund, 13 août 2014*

e	*laps-e-ha*	*e*	*tikn-e-ha*
ART.DEF.M.SG.OBL	crayon-M.SG.OBL-INSTR	ART.DEF.M.SG.OBL	autre-M.SG.OBL-INSTR

« avec le petit crayon »

(240) *F24AČ, Korçë, 16 août 2014*

o	*bišti/*	*e*	*sir-a-k-o*	*e*	*bar-a-k-o*
ART.DEF.M.SG	queue	ART.DEF.M.SG.OBL	ail-F.SG.OBL-GEN-M.SG	ART.DEF.M.SG.OBL	grand-F.SG.OBL-GEN-M.SG

« la queue de la grande tête d'ail [litt. du grand ail] »

On se souvient que dans les syntagmes monodéfinis, l'adjectif postposé est sous-spécifié. Ici il est sur-spécifié, car il porte les informations de genre et de nombre, mais aussi de cas, avec des affixes nominaux et non adjectivaux. La définitude apparaît presque comme un trait supplémentaire porté par l'adjectif, à ceci près qu'elle est exprimée par un morphème libre et non par un affixe.

[1] Éventuellement situé après l'adjectif de taille, dans l'exemple (233).

2.3.2.4 Répétition de la préposition

On peut distinguer deux types de prépositions : celles qui restent distinctes de l'article défini et celles qui fusionnent avec lui (clitique). Lorsque la préposition est distincte de l'adjectif, il ne se produit rien de particulier :

(241) *F47M, Llakatund, 13 août 2014*

paše	*i*	*sir/*	*i*	*bar-i*
près	ART.DEF.F.SG	ail(F.SG)	ART.DEF.F.SG	grand-F.SG
PREP	**DEF**	**N/**	**DEF**	**A**

« à côté du grand ail »

(242) *M19M(2), Llakatund, 13 août 2014*

ka	*o*	*lapsi*	*o/*	*i*	*vərð*
à	ART.DEF.M.SG	crayon(M.SG)	ART.DEF.M.SG	ART.DEF.M.SG$_{albanais}$	jaune
PREP	**DEF**	**N/**	**DEF**	**A**	**PREP**

« au crayon/ jaune »

En (242), l'adjectif est emprunté à l'albanais : l'emprunt est lexical (le lexème *vərð*), morpho-syntaxique (l'adjectif vient avec son article de connexion, *i* au lieu de *o* qui vient d'être prononcé) et syntaxique (l'adjectif est postposé au nom). Nous y reviendrons en section 2.4, p. 329.

Lorsque la préposition et l'article ont fusionné, on peut trouver un syntagme polydéfini classique :

(243) *F23AČ, Voskopojë, 28 juillet 2014*

to	*manaćyri*	*o*	*lol-o*
à.ART.DEF.M.SG	vernis_à_ongles(M.SG)	ART.DEF.M.SG	rouge-M.SG
PREP.DEF	**N**	**DEF**	**A**

« au vernis à ongles rouge »

(244) *M19M(2), Llakatund, 13 août 2014*

ki	*ana*	*i*	*vaver*
à.ART.DEF.F.SG	côté(F.SG)	ART.DEF.F.SG	autre
PREP.DEF	**N**	**DEF**	**A**

« de l'autre côté »

(245) *M18Č, Korçë, 24 juillet 2014*

paše	*to*	*stilolapsi*	*o*	*lol-o*
près	à.ART.DEF.M.SG	stylo(M.SG)	ART.DEF.M.SG	rouge-M.SG
PREP	**PREP.DEF**	**N**	**DEF**	**A**

« près du stylo rouge »

(246) *M19M(2), Llakatund, 13 août 2014*

opre	*ki*	*patata*	*i*	*vaver*
en_haut	à.ART.DEF.F.SG	pomme_de_terre(F.SG)	ART.DEF.F.SG	autre
PREP	**PREP.DEF**	**N**	**DEF**	**A**

« en haut près de l'autre pomme de terre »

Lorsque l'article a fusionné avec la préposition, l'ensemble fusionné peut être répété. On a alors une structure assez lourde : PREP.DEF N PREP.DEF A.

(247) *F22M, Llakatund, 13 août 2014*

ki	*ana*	*ki*	*vaver*
à.ART.DEF.F.SG	côté(F.SG)	à.ART.DEF.F.SG	autre
PREP.DEF	**N**	**PREP.DEF**	**A**

« de l'autre côté »

(248) *F44AČ, Korçë, 23 juillet 2014*

ti	*patata*	*ti*	*bar-i*
à.ART.DEF.F.SG	pomme_de_terre(F.SG)	à.ART.DEF.F.SG	grand-F.SG
PREP.DEF	**N**	**PREP.DEF**	**A**

« à la grande pomme de terre »

Ce fait peut être interprété comme une preuve supplémentaire du haut degré de fusion de la préposition et de l'article défini (Boretzky 2000b : 44). Il est intéressant de noter qu'en présence de deux adjectifs qui encadrent le nom, il n'y a pas nécessairement trois articles définis :

(249) *F65A, Korçë, 1er août 2014*

ti	*xurd-i*	*patata/*	*ti/*	*vaver*
à.ART.DEF.F.SG	tout_petit-F.SG	pomme_de_terre(F.SG)	à.ART.DEF.F.SG	autre
PREP.DEF	**A**	**N**	**PREP.DEF/**	**A**

« à l'autre/ toute petite/ pomme de terre »

On n'a donc pas [PREP.]DEF A <u>[PREP.]DEF</u> N [PREP.]DEF A, mais [PREP.]DEF A N [PREP.]DEF A.

La structure se répète autant de fois qu'il y a d'adjectifs : PREP.DEF N PREP.DEF A PREP.DEF A :

(250) *F44AČ, Korçë, 23 juillet 2014*

ti	*domata*	*ti*	*lol-i/*	*ti*	*bar-i/*	*ti*	*lol-i/*
à.ART.DEF.F.SG	tomate(F.SG)	à.ART.DEF.F.SG	rouge-F.SG	à.ART.DEF.F.SG	grand-F.SG	à.ART.DEF.F.SG	rouge-F.SG
PREP.DEF	**N**	**PREP.DEF**	**A/**	**PREP.DEF**	**A/**	**PREP.DEF**	**A/**

« à la tomate rouge/ grande/ rouge »

(251) *F65A, Korçë, 1er août 2014*

ti	*patata*	*ti*	*tikn-i/*	*ti*	*xurd-i*
à.ART.DEF.F.SG	pomme_de_terre(F.SG)	à.ART.DEF.F.SG	petit-F.SG	à.ART.DEF.F.SG	tout_petit-F.SG
PREP.DEF	**N**	**PREP.DEF**	**A/**	**PREP.DEF**	**A**

« à la petite tomate, toute petite »

L'exemple (250) est prononcé juste avant l'exemple (227). Au référent est d'abord assignée une propriété inutile à l'identification (une tomate est bien souvent rouge à Korçë – du moins toutes les tomates sur la table sont-elles rouges). Un tel syntagme est du même type que celui des cobras venimeux de (Kolliakou 2004 : 273) : l'autrice explique alors qu'un syntagme polydéfini est impossible avec la combinaison *cobras+venimeux*, car « *pragmatically non-restrictive terms cannot be polydefinites* ». En effet, la réalité ne laisse pas d'autre interprétation possible que hiérarchique, où la totalité des cobras sont inclus dans les entités venimeuses : il n'y a pas d'interprétation conjointe possible, c'est-à-dire pas de restriction pragmatique possible. De même, toutes les tomates de la table sont incluses dans les entités rouges, d'où le rire des personnes présentes et de la locutrice elle-même. Le syntagme est grammaticalement possible, mais il est inutile sur le plan informatif et pragmatique. C'est pourquoi la locutrice ajoute un autre adjectif, cette fois de taille, en conservant le patron polydéfini qui vient d'être utilisé. On peut qualifier cet ajout d'auto-correction. Notons que suite à ce syntagme et à cette auto-correction, le participant B, M13AČ, (et moi-même) avons ri, ce qui pourrait expliquer l'ajout ultime de l'adjectif « inutile » *loli*, soit par agacement, soit par autodérision de la locutrice A, F44AČ.

Comment expliquer qu'ait pu être prononcé un syntagme qualifié d'*infelicitous* par la littérature sur le grec moderne ? Lorsque la locutrice F44AČ prononce l'exemple (227), elle a déjà mentionné le référent en question : elle a commencé l'enregistrement en saisissant cet objet (la grosse tomate, DEF N DEF A). Elle a manipulé ensuite divers objets (pommes de terre, vernis à ongles, crayon). Elle prend ensuite « la tomate » (DEF N), sans plus d'information, pour la mettre au fond. Le participant B, M13AČ, devine qu'il s'agit d'une autre tomate et non de la grosse qui vient d'être mentionnée. Il vérifie cette intuition en demandant : *i uštime?*, « la pointue ? », ce qui est confirmé par la locutrice A. Celle-ci prend ensuite *jek moliði*, un crayon, là aussi sans plus d'information, ce qui là aussi pousse B à demander : *so/ molivi?*, « quel/ crayon ? »... mais la locutrice A, F44AČ, est déjà en train de poursuivre sa phrase : « on va la mettre à/ à/ côté de la pomme de terre » (en employant le pronom personnel objet féminin *la*, la, pour désigner en fait *moliði*, crayon, de genre masculin, ce qui peut indiquer qu'elle est déjà en train de penser à la tomate, de genre féminin). M13AČ doit alors l'interrompre pour répéter sa question : *s/ so/ nđyra/ so/ XXX*[1]*?*, « qu/ quelle/ couleur/ quel/ XXX ? ». La locutrice A dit alors deux fois *o kalo*, le noir (DEF A). C'est à ce moment qu'elle prononce *ka čhuvas la/ ti domata ti loli/ ti bari/ ti loli*, « on va la mettre/ à côté de la tomate rouge/ grande/ rouge » – toujours avec une confusion de genre sur le pronom personnel. On peut donc interpréter l'irruption d'un adjectif de couleur pour qualifier la tomate, qui nécessite un adjectif de taille, comme une confusion avec le dernier référent mentionné, le crayon, qui nécessitait, lui, un adjectif de couleur.

[1] Trois syllabes inintelligibles car couvertes par la voix de la participante A.

Lorsque la préposition se compose de plusieurs éléments, seule est répétée la partie (monosyllabique) à laquelle est cliticisé l'article – qu'il y ait un adjectif (252) ou plusieurs (253), (254).

(252) *F27A, Korçë, 21 juillet 2014*

paš	*ti*	*domatka*	*ti*	*bar-i*
près	à.ART.DEF.F.SG	tomate(F.SG)	à.ART.DEF.F.SG	grand-F.SG
PREP	**PREP.DEF**	**N**	**PREP.DEF**	**A**

« à côté de la grande tomate »

(253) *F16AČ, Korçë, 7 août 2014*

paš	*to*	*manyćyri*	*to*	*vaver*	*to*	*parn-o*
près	à.ART.DEF.M.SG	vernis_à_ongles(M.SG)	à.ART.DEF.M.SG	autre	à.ART.DEF.M.SG	blanc-M.SG
PREP	**PREP.DEF**	**N**	**PREP.DEF**	**A**	**PREP.DEF**	**A**

« à côté de l'autre vernis à ongles blanc »

(254) *F16AČ, Korçë, 7 août 2014*

paš	*ti*	*kultulka*	*ti*	*tikn-i/*	*ti*	*vaver*
près	à.ART.DEF.F.SG	pomme_de_terre(F.SG)	à.ART.DEF.F.SG	petit-F.SG	à.ART.DEF.F.SG	autre
PREP	**PREP.DEF**	**N**	**PREP.DEF**	**A/**	**PREP.DEF**	**A**

« à côté de la petite tomate… l'autre »

On a ici une confirmation de l'importance de la place des adjectifs lorsqu'ils sont postposés : si tous sont contrastifs, l'information principale permettant l'identification est placée en première position. *Vaver* suffit à contraster en (253) puisque le référent entre en opposition binaire avec un autre, mais il ne suffit pas en (254), où le référent se trouve en opposition ternaire, d'où son rôle secondaire[1].

Notons ici deux cas de répétition de la préposition, où ce qui joue le rôle de l'adjectif est en réalité un nom au génitif privé de son article défini :

(255) *F22M, Llakatund, 13 août 2014*

ko/	*ki*	*ana*	*ki*	*udar-es-Ø-i*
à.ART.DEF.M.SG	à.ART.DEF.F.SG	côté(F.SG)	à.ART.DEF.F.SG	porte-M.SG.OBL-GEN-F.SG
PREP.DEF/	**PREP.DEF**	**N**	**PREP.DEF**	**SGEN**

« de/ du côté de la porte »

(256) *F22M, Llakatund, 13 août 2014*

ko	*krahu*	*ki*	*pendžer-ija-k-o*
à.ART.DEF.M.SG	bras(M.SG)	à.ART.DEF.F.SG	fenêtre-F.SG.OBL-GEN-M.SG
PREP.DEF	**N**	**PREP.DEF**	**SGEN**

« de/ du côté de la fenêtre »

[1] L'exemple (252) fait partie de la phrase énigmatique citée plus haut dans l'analyse de l'exemple (233).

Ce fait inattendu peut être interprété comme une preuve de la nature de modifieur adnominal du syntagme nominal en romani – qui joue donc le même rôle qu'un adjectif. Boretzky (2000b : 42) écrit que les cas d'adjectivisation sont rares et qu'il s'agit de cas particuliers dans des variétés particulières. L'emploi du nom au génitif en lieu et place de l'adjectif au sein d'une structure polydéfinie, dans les exemples (255) et (256), montre la souplesse et la productivité tant du groupe génitival à statut adjectival que de la polydéfinitude.

2.3.2.5 Témoins

Les témoins ne produisent qu'un seul syntagme polydéfini (257). Cette unique occurrence est en fait de structure DEF N/ DEF A, avec une pause entre les deux sous-parties du syntagme. Mais le deuxième article défini n'est pas le même que le premier :

(257) *M27Dž, Ohrid, 30 juillet 2014*

o	*lapsi/*	*to*	*portokal*
ART.DEF.M.SG	crayon(M.SG)	ART.DEF.M.SG	orange
DEF	**N/**	**DEF**	**A**

« le crayon/ orange »

Doit-on en déduire que le locuteur utilise un article *to* en romani ? Il peut s'agir d'un même article défini supplémentaire que les Roms de cette région centrale des Balkans auraient ajouté aux deux articles romani communs *o* et *i*. Il est également possible que le *to* de Korçë et le *to* du témoin soient d'origine différente : grecque (article défini neutre το, *to*) chez les Korçiens où il est en variation libre avec l'article défini masculin hérité *o* pour déterminer des noms récemment empruntés (à l'albanais et au grec) ; macédonienne (article défini neutre то, *to*, qui est cependant clitique et postposé) chez le Macédonien, en variation libre non encore déterminée.

On trouve en revanche chez le témoin originaire du Kosovo les deux syntagmes suivants :

(258) *M43X, Korçë, 20 juillet 2014*

o	*špicast-o*	*patlidžano*	*o*	*lol-o*
ART.DEF.M.SG	pointu-M.SG	aubergine(M.SG)	ART.DEF.M.SG	orange-M.SG
DEF	**N**	**A**	**DEF**	**A**

« la tomate pointue » (2 occurrences)

(259) *M43X, Korçë, 20 juillet 2014*

o	*bar-o*	*patlidžano/*	*o*	*lol-o*
ART.DEF.M.SG	grand-M.SG	aubergine(M.SG)	ART.DEF.M.SG	orange-M.SG
DEF	**A**	**N/**	**DEF**	**A**

« la tomate/ grande »

Dans la variété de son groupe (xoraxane), *tomate* se dit *aubergine rouge*. Lors d'une discussion sur les différences lexicales avec son interlocuteur albanais, il oppose la tomate, *patlidžano lolo*, aubergine rouge, et l'aubergine, *patlidžano kalo*, *aubergine noire*, au moyen de syntagmes monodéfinis à adjectif postposé DEF N A. Lorsqu'il mentionne simplement l'existence de ces vocables, l'adjectif est antéposé, A N (« *ta pal si amen kalo patlidžano*$_1$/ *bordžin*$_2$ *kaj phenel ose/ patlidžano/ kalo*$_1$ », « et on a aussi l'aubergine noire$_1$/ aubergine$_2$ elle s'appelle ou encore/ aubergine/ noire$_1$ »).

Ici, dans un contexte où la tomate ne contraste pas avec l'aubergine, c'est-à-dire un contexte où ce n'est pas la couleur qui est contrastive, l'adjectif qui l'exprime est postposé au nom et accompagné d'un article. Le contraste entre les tomates réside dans leur taille (*grande*, *petite*) et dans leur aspect (*pointue*), et l'adjectif qui exprime ces propriétés est antéposé au nom en (258) et (258), ainsi que dans un troisième syntagme *o cikno patlidžano*, la petite aubergine [tomate], de forme DEF A N. On voit que dans cette variété, la fonction de contraste est assumée par la position antéposée au nom, contrairement à ce que l'on a pu observer dans les variétés albanaises. La fonction de qualification (en réalité, le simple nom du référent), qui n'est pas ici contrastive, est ici assumée par la polydéfinitude. Ce syntagme comprend donc deux adjectifs qui ont une syntaxe et des rôles différents : c'est pourquoi nous avons choisi de l'intégrer dans la catégorie des syntagmes complexes plutôt que dans celle des syntagmes polydéfinis.

2.3.3 Syntagmes introduits par un autre déterminant

2.3.3.1 Tests 1 et 2

Nous avons vu jusqu'à présent où était placé l'adjectif dans les syntagmes introduits par un article défini. Qu'en est-il de ceux introduits par un autre déterminant ?

Tableau 68 : Position de l'adjectif dans les autres syntagmes produits par les participants des tests 1 et 2

	DEF N DEF A	A postposé	A antéposé	complexes	Total
Syntagmes déterminés par DEM	1 (5 %)	12 (60 %)	3 (15 %)	4 (20 %)	20 (100 %)
Syntagmes déterminés par POSS	0 (0 %)	2 (100 %)	0 (0 %)	0 (0 %)	2 (100 %)
Syntagmes déterminés par IND	0 (0 %)	36 (61,02 %)	18 (30,51 %)	5 (8,47 %)	59 (100 %)
Syntagmes déterminés par QUANT / NUM	0 (0 %)	1 (50 %)	0 (0 %)	1 (50 %)	2 (100 %)
Syntagmes déterminés par Ø	-	10 (43,48 %)	11 (47,83 %)	2 (8,69 %)	23 (100 %)
Total	1 (0,94 %)	61 (57,55 %)	32 (30,19 %)	12 (11,32 %)	106 (100 %)

Dans les syntagmes déterminés par un autre déterminant, la tendance est la même que dans les syntagmes définis : l'adjectif y est en grande majorité postposé (57,55 % + 0,94 %) plutôt qu'antéposé (30,19 %).

2.3.3.2 Témoins

Chez les témoins, l'adjectif était antéposé de manière quasi systématique dans les syntagmes introduits par un article défini. Qu'en est-il de ceux introduits par un autre déterminant ?

Tableau 69 : Position de l'adjectif dans les autres syntagmes produits par les témoins

	DEF N DEF A	A postposé	A antéposé	complexes	Total
Syntagmes déterminés par DEM	0 (0 %)	0 (0 %)	2 (100 %)	0 (20 %)	20 (100 %)
Syntagmes déterminés par POSS	0 (0 %)	0 (0 %)	2 (100 %)	0 (20 %)	20 (100 %)
Syntagmes déterminés par IND	0 (0 %)	1 (33,33 %)	1 (33,33 %)	1 (33,33 %)	3 (99,99 %)
Syntagmes déterminés par QUANT / NUM	0 (0 %)	0 (0 %)	0 (0 %)	0 (0 %)	0 (0 %)
Syntagmes déterminés par Ø	-	0 (0 %)	8 (100 %)	0 (0 %)	8 (100 %)
Total	0 (0 %)	1 (6,67 %)	13 (86,67 %)	1 (6,67 %)	15 (100,01 %)

Dans les syntagmes déterminés par un autre déterminant, la tendance est la même que dans les syntagmes définis : l'adjectif y est toujours antéposé, de manière presque catégorique (86,67 %).

2.3.4 Synthèse

Dans le corpus des locuteurs albanais, l'adjectif est en majorité postposé dans tous les syntagmes (monodéfinis, polydéfinis, introduits par un autre déterminant). On peut donc dire qu'en romani d'Albanie, dans le contexte énonciatif des tests, la postposition est la position non-marquée de l'adjectif, contrairement à l'antéposition qui est marquée. La syntaxe est ici tout à fait différente de ce qui est décrit dans la littérature existante sur le romani.

La polydéfinitude se construit toujours, à deux exceptions près (DEM DEF A DEF N et DEF A/ DEF N), avec un adjectif postposé (DEF N DEF A). C'est une différence importante avec le grec qui autorise à la fois DEF A DEF N et DEF N DEF A. Ainsi, contrairement au grec (qui autorise DEF A N mais pas *DEF N A), l'adjectif romani est plus flexible dans les syntagmes monodéfinis (il autorise DEF A N et DEF N A) que polydéfinis. Lorsque le syntagme présente plusieurs adjectifs, ils peuvent soit encadrer le nom dans les syntagmes monodéfinis (DEF A N A) et dans les syntagmes polydéfinis (DEF A N DEF A), soit

lui succéder dans des syntagmes partiellement polydéfinis (DEF N DEF A/ A et DEF N A/ DEF A) ou totalement polydéfinis (DEF N DEF A DEF A). Nous avons constaté que lorsque le syntagme présente plusieurs adjectifs qui encadrent le nom, celui antéposé transmet une propriété secondaire, peu contrastive ou de moindre importance, tandis que celui postposé transmet une propriété contrastive permettant l'identification, donc de première importance. Lorsque les adjectifs sont postposés au nom, celui qui vient en premier est celui de première importance, qui permet l'identification, tandis que celui qui vient en second est de moindre importance. Nous avons constaté une exception (exemple (226) où une locutrice (F, +60 ans, 0-5 années d'éducation, non-militante, arli) était par trois fois réticente à postposer l'adjectif permettant l'identification, alors que cela semblait attendu.

Dans le corpus des témoins, l'adjectif est quasiment toujours antéposé (monodéfinis, polydéfinis, introduits par un autre déterminant). On note une seule occurrence de polydéfinitude, et trois occurrences de syntagmes complexes DEF A N(/) DEF A, où l'adjectif qualificatif « simplement » informatif est postposé, tandis que l'adjectif contrastif permettant l'identification est antéposé – ce qui est le contraire de la syntaxe des locuteurs albanais.

Nous avons constaté chez les locuteurs albanais une variation de la flexion adjectivale. Nous suivons la stratification de la flexion nominale établie par Matras (2002 : 78-94), qui distingue un *layer I* (*nominative and oblique*) de nature fusionnelle, un layer II agglutinante (*dative, locative, ablative, instrumental, genitive*) de nature, un layer III (*analytic adpositions*) de nature prépositionnelle, ainsi que le vocatif. Le Tableau 70 représente les différentes combinaisons possibles dans le syntagme nominal défini modifié par un adjectif, en fonction de la morphologie de l'adjectif et du nom (concordantes ou non), de la syntaxe adjectivale (adjectif antéposé ou postposé au nom), et de la mono- ou polydéfinitude. L'existence d'une structure est notée par ✓, sa possible existence par ? et son impossibilité par *. Les expressions présentes dans notre corpus sont mises en gras.

On s'attend à A1 (*o lolo lapsi*, le crayon rouge), D1 (*to lolo lapsi*, au crayon rouge), B2 (*e tikne manušes*, le petit homme), C2 (*e lače vastesa*). L'existence de C3 (*e lapseha e tikneha*, avec le petit crayon) est mentionnée comme possible par la littérature. En revanche, l'existence de C1 (*e lapsesa lolo*, avec le crayon rouge) et de D4 (*ti kultulka ti bari*, à la grosse pomme de terre) est tout à fait inattendue.

Tableau 70 : Economie du syntagme nominal dans les tests 1 et 2 : morphologie et syntaxe

		A	B	C	D
		Nom : strate I cas direct	Nom : strate I cas oblique	Nom : strate II	Nom : strate III
1	Adjectif : strate I cas direct	✓	* & ?	* & ✓	✓
		o lolo lapsi ***o lapsi lolo*** ***o lapsi o lolo*** ***o roza o manyćyri***	**e tikno manušes* *?e manušes tikno* **e manušes e tikno*	**e lolo lapsesa* ***e lapsesa lolo*** **e lapsesa e lolo*	***to lolo lapsi*** ***to lapsi lolo*** ***ko manyćyri o parno***
2	Adjectif : strate I cas oblique	*	✓ & ?	✓ & ?	*
		**o tikne lapsi* **o lapsi tikne* **o lapsi o tikne*	*e tikne manušes* *?e manušes tikne* *?e manušes e tikne*	***e lače vastesa*** ***témoin*** *?e vastesa lače* *?e vastesa e lače*	**to tikne lapsi* **to lapsi tikne* **to lapsi o tikne*
3	Adjectif : strate II	*	*	? & ✓	*
		**o lolesa lapsi* **o lapsi lolesa* **o lapsi o lolesa*	**e tiknesa manušes* **e manušes tiknesa*	*?e lolesa lapsesa* *?e lapsesa lolesa* **e lapseha e tikneha**	**to lolesa lapsi* **to lapsi lolesa*
4	Adjectif : strate III (si polydéfinitude)	*	*	*	? & ✓
		[l'adjectif ne peut pas être introduit par une préposition si le nom ne l'est pas déjà lui-même]			?ti bari ti kultulka **ti kultulka ti bari**

2.4 Variable 3 : TYPE D'ADJECTIF

Quels types d'adjectifs peuvent entrer dans un syntagme polydéfini ? Nous entendons ici par « type »[1] deux aspects différents : l'origine et le type de l'adjectif ainsi que sa combinaison avec le nom.

Est-ce qu'un lexème emprunté ou traité comme un emprunt pourrait bloquer la polydéfinitude ? Cette question n'a pas été posée pour le grec moderne. Matras (2002 : 209-213) montre que l'emprunt grammatical est de trois types en romani : « *compartmentalisation* », **compartimentation** (scission de la grammaire entre une morphologie « héritée » et une « empruntée »), **convergence** (adaptation des ressources héritées à un modèle extérieur, notamment au niveau syntaxique) et **fusion** (absence de séparation pour certaines catégories, voire pour certaines fonctions linguistiques).

Adamou et Granqvist (2015) confirment pour le romani thrace de Grèce et pour le romani finnois que la morphologie est compartimentée : les verbes d'origine turque en romani thrace sont conjugués au moyen d'affixes turcs, et les verbes d'origine finnoise en romani finnois sont conjugués au moyen d'affixes finnois (affixes de temps, de mode, d'aspect et de personne) – les phonèmes des langues d'emprunt étant conservés. Il en va de même pour la flexion nominale en romani finnois – mais pas en romani thrace, où les noms turcs suivent la morphologie du romani (affixes de cas et de genre) (Adamou et Granqvist 2015). Il s'agit de variétés de romani particulières, catégorisées comme des *unevenly mixed languages*, ce qui n'est pas le cas des variétés romani d'Albanie. Nous nous demanderons comment est traité l'adjectif épithète : conserve-t-il la morphologie de sa langue d'origine (cf. les noms finnois en romani finnois) ou non (cf. les noms turcs en romani thrace) ?

La syntaxe quant à elle semble grandement absente des études sur l'emprunt lexical : la synthèse de Winford (2003) par exemple étudie l'intégration des emprunts sous l'angle de la phonologie (section 2.7.1) et de la morphologie (section 2.7.2), et les conséquences structurelles de l'emprunt lexical sous l'angle de la phonologie (section 2.9.1), de la morphologie (section 2.9.2) et du lexique (section 2.9.3). Matras (2009), dans son chapitre sur l'emprunt lexical, consacre cependant une section à l'emprunt des adjectifs et des adverbes (section 7.4) : il s'y penche entre autres sur la morphologie et la syntaxe. Ses nombreux exemples de langues sont résumés dans le Tableau 71.

Tableau 71 : Typologie de l'emprunt adjectival chez Matras (2009)

	Syntaxe	Morphologie
allemand	✓: antéposition	✓: flexion héritée
hébreu	✓: postposition	X : assignation à une classe flexionnelle particulière
yiddish	✓: antéposition	✓: flexion héritée (adaptée si besoin)

[1] Cette variable a été écartée, pour des raisons de place, dans les analyses présentées dans Tirard (2017).

	Syntaxe	Morphologie
romani	? ✓: antéposition (Burgenland)	✓: flexion héritée (adaptée si besoin) X : absence de morphologie (Burgenland)
persan, kurde	✓: postposition	✓: absence de flexion
turc	✓: antéposition	✓: absence de flexion
ourdou	?	X : absence de morphologie
anglais	X : postposition	X : flexion de la langue d'origine
maltais	✓: postposition (héritée & empruntée)	✓: flexion héritée ou partiellement empruntée X : absence de flexion

Nous allons voir si, en romani d'Albanie, les adjectifs tendent à être intégrés syntaxiquement dans la position de l'épithète, c'est-à-dire DEF A N, ou s'ils sont employés dans la position de la langue d'origine (convergence) lorsque celle-ci est différente. Nous rechercherons notamment si des adjectifs empruntés peuvent entrer dans un syntagme polydéfini. Il s'agira également de placer le romani d'Albanie dans la typologie établie par Matras (2009).

Enfin, nous nous demanderons si tous les types d'adjectifs peuvent entrer dans un syntagme monodéfini, en fonction de leur sémantique (taille, aspect, couleur, autres sens) et en fonction de la manière dont ils se combinent avec le nom. La littérature sur le grec moderne nous apprend en effet que seuls les adjectifs amenant une interprétation restrictive peuvent entrer dans un syntagme polydéfini (cf. section 2.4.2, p. 179).

Remarques sur l'étymologie

Les différentes variétés albanaises présentent une histoire linguistique différente, donc des emprunts lexicaux et grammaticaux différents à des langues différentes. C'est pourquoi nous ne considérerons comme hérité que ce qui relève du socle commun de toutes les variétés de romani, soit le vocabulaire indo-aryen (*Proto-Romani,* Matras 2002 : 18-20) et la strate lexicale empruntée aux premières langues de contact avant l'arrivée en Europe (persan, arménien, langues caucasiennes, grec « byzantin »), avant la dispersion supposée des groupes de Roms à travers l'Europe à l'époque byzantine tardive (*Early Romani,* Matras 2002 : 18-20). Il est difficile de dater les emprunts postérieurs à cette époque (mots turcs en arli, mots romans en čergar I, etc.). En outre il reste à étudier si les variétés albanaises partagent des emprunts autres que ceux, récents, à l'albanais, et déterminer lesquels. On pourrait en revanche stratifier, variété par variété, les emprunts en suivant Matras (2002 : 195) (cf. section 2.2.1, p. 34). On peut faire l'hypothèse suivante pour l'arli :

- des emprunts à l'albanais comme langue de contact principale : *verð, jaune* (albanais *verdh*, 'jaune'),
- des emprunts au grec comme langue de contact secondaire : *ble, bleu* (grec μπλε, *ble*, 'bleu'),
- des emprunts au turc comme ancienne langue de contact : *saria, jaune* (turc *sarı*, jaune).

La décision de considérer les lexèmes empruntés à la *recent L2* et à la *older L2* comme des emprunts et non comme des lexèmes hérités s'explique par un positionnement au niveau du romani d'Albanie global, et non au niveau de telle ou telle variété. Cela peut se discuter, dans la mesure où de nombreux emprunts au grec, au roman, au slave et au turc sont en fait ressentis comme « hérités » par les locuteurs. Ainsi *saria* (arli) et *žolto*, jaune, emprunté au slave du sud par le mečkar, sont-ils considérés par les locuteurs comme du vocabulaire fondamental de leur variété voire du romani en général – selon leur familiarité plus ou moins importante avec les autres variétés. Ce ressenti provient du fait qu'un emprunt à une langue autre que l'albanais, langue dominante commune, est souvent un emprunt ancien, se distinguant peu du reste du lexique. Comme ces langues d'emprunt ne sont plus, actuellement, parlées par une majorité de locuteurs (cf. section 2.2.1, p. 34), l'origine des lexèmes est devenue opaque, ainsi que leur signification pour les locuteurs d'autres variétés. C'est pour cette raison que de tels adjectifs sont ressentis comme hérités par les locuteurs. Une même personne pouvait ainsi, au cours d'un même enregistrement, exprimer la couleur jaune par *verð*, emprunté à l'albanais, puis se reprendre et utiliser *saria*, considéré comme « meilleur » car hérité (cf. Adamou et Arvaniti 2014 : 226) qui décrivent des réflexes lexicaux similaires en romani thrace de Grèce) – alors même que *saria* est lui aussi un emprunt. Nous considérerons en revanche que les emprunts spécifiques à une variété font partie de son stock lexical propre, mais non du stock lexical commun du romani d'Albanie.

2.4.1 Poids des adjectifs et type syntaxique

On a vu en section 2.2.2, p. 74, que le poids respectif du nom et de l'adjectif jouerait un rôle : l'élément le plus lourd tend à la postposition. Qu'en est-il dans notre corpus ?

Tableau 72 : Poids et place de l'adjectif : quelques exemples du corpus semi-spontané

Adjectifs	DEF A N	DEF N A	DEF N DEF A
1 syllabe	*o ros mandalàti* la pince à linge rose	*o manyćyri roz* le vernis à ongles rose	*o kalemi o ros* le crayon rose
2 syllabes	*i bari patata* la grande pomme de terre	*i patata bari* la grande pomme de terre	*i patata i bari* la grande pomme de terre
3 syllabes	*o saria lapsi* le crayon jaune	*o lapsi saria* le crayon jaune	*o lapsi o saria* le crayon jaune
4 syllabes	*ti maškaruni kutulka* à la pomme de terre moyenne	*o lapsi portokali* le crayon orange	*to lapsi to/ portokali* au crayon/ orange
5 syllabes		*o lapsi portokalia* le crayon orange	*o lapsi o portokalia* le crayon orange

Le poids de l'adjectif ne semble pas un critère pertinent pour décider de la position de l'adjectif dans ce corpus.

Tableau 73 : Poids et place du nom : quelques exemples du corpus semi-spontané

Noms	DEF A N	DEF N A	DEF N DEF A
1 syllabe	*i bari sir* le grand ail	*ti sir bari* au grand ail	*i sir i bari* le grand ail
2 syllabes	*i vaver kapsa* l'autre pince à linge	*i kapsa roza* la pince à linge rose	*i kapsa i roza* la pince à linge rose
3 syllabes	*o lolo kalemi* le crayon rouge	*o kalemi lolo* le crayon rouge	*o kalemi o lolo* le crayon rouge
4 syllabes	*o parno manyćyri* le vernis à ongles blanc	*o manyćyri parno* le vernis à ongles blanc	*o manyćyri o parno* le vernis à ongles blanc

Le poids du nom ne semble pas non plus pas un critère pertinent pour décider de la position de l'adjectif dans ce corpus, puisque l'on trouve des noms de toutes tailles en toute place.

Nous avons vu en section 2.3.2, p. 164, que les adjectifs des syntagmes polydéfinis doivent être prédicatifs, et pouvoir exercer la fonction d'attribut du sujet. Les adjectifs de notre corpus sont tous prédicatifs, à l'exception de *vaver*, autre. Il en va de même en grec, où l'adjectif άλλος, *allos*, peut lui aussi entrer dans une structure polydéfinie (Lekakou et Szendrői 2012 : 133).

2.4.2 Origine et sémantique des adjectifs dans les syntagmes polydéfinis

Nous avons classé les adjectifs entrant en jeu dans les syntagmes polydéfinis en fonction de leur sémantique et de leur étymologie. La forme phonétique des adjectifs s'est révélée extrêmement variable : nous n'en présentons ici qu'une par adjectif, au masculin singulier car la désinence *-o* qui la caractérise ne donne pas lieu à palatalisation, contrairement au féminin singulier *-i* et au pluriel *-e*. Le Tableau 76 montre que tous les adjectifs sont hérités, à l'exception des adjectifs de couleur et *mesme*, moyen (adjectif de taille emprunté à l'albanais). Cela peut être une caractéristique de ces variétés, ou bien une conséquence du test, où les locuteurs se sentent évalués sur leur maîtrise de la langue, et en particulier sur leur connaissance du lexique hérité (notamment lorsque je répondais qu'il s'agissait du but de la tâche, cf. section 1.1.2.1, p. 289).

Au contraire, les adjectifs de couleur sont tous empruntés, à l'exception de rouge, blanc et noir. Aucune variété du romani n'a conservé d'autre couleur que celles-ci. Elles semblent en effet plus « fondamentales » que les autres, si l'on en croit Matras (2009 : 187) qui cite Berlin et Kay (1969). Ces derniers ont établi une hiérarchie de différentiation des termes de couleur : noir & blanc > rouge > jaune & vert > bleu. Cette hiérarchie prédit que noir et blanc sont les couleurs les plus susceptibles d'être représentées, dans les langues du monde, par un terme ou une expression indépendante, au contraire de bleu qui existe dans le moins de langues du monde. Noir et blanc sont donc des termes de couleur « universelles », suivies par le rouge. En romani, la probabilité d'emprunt est inversement proportionnelle à la position qu'occupe le terme de couleur dans la hiérarchie. C'est pourquoi jaune, vert, bleu (mais aussi orange, rose…) sont toujours empruntés en romani : ils occupent le bas de l'échelle hiérarchique. Le Tableau 75 montre que, lexicalement, les adjectifs de couleur sont à

76,92 % empruntés (10 adjectifs sur 13) et trois d'entre eux hérités (23,08 %). En revanche ces derniers représentent la majorité des occurrences employées dans le corpus polydéfini (106 adjectifs de couleur sur un total de 195, soit 54,64 %).

Sous quelle forme se présentent les adjectifs empruntés ?

Tableau 74 : Morphologie des adjectifs empruntés dans les syntagmes polydéfinis des tests 1 et 2 (témoins exclus)

Formes	Traduction	Morphologie
blu	bleu	M.SG
bluja		
ble		
gri	gris	M.SG
grija		
ješil	vert	M.SG
ješili		M.SG, F.SG
ješilka		F.SG
portokali	orange	M.SG
portokalija		

Formes	Traduction	Morphologie
ros	rose	M.SG, F.SG
roza		
roz		F.SG
sarija	jaune	M.SG
i/e verð	jaune	M.SG, F.SG
i vərði		M.SG
višnja	cerise	M.SG
žolto	jaune	M.SG, F.SG
e mesme	moyen	F.SG

Le corpus polydéfini ne permet pas de tester la différence singulier ~ pluriel, ni la différence cas direct ~ autres cas. On constate avec le Tableau 74 que certains adjectifs ne sont employés que dans une configuration, par exemple *sarija*, employé avec *lapsi*, *moliði*, *kalemi*, crayon, et *stilolapsi*, stylo, qui sont tous de genre masculin. D'autres adjectifs tels que *višnja* ne sont employés qu'une fois, ce qui ne permet pas d'établir une comparaison.

Les adjectifs qui sont employés dans une plus grande variété d'occurrences présentent une forme invariable en genre, par exemple *ješili, ros, verð, žolto*. L'intégration à la morphologie romani d'un adjectif slave tel que *žolto* aurait dû donner **žolti* au féminin singulier ; le respect de la morphologie de la langue d'origine aurait dû donner **žolta*. Difficile de juger pour *ješili*, *verð* et *ros*, qui ne varieraient pas en genre en albanais, au nominatif singulier déterminé. Notons que deux adjectifs empruntés à l'albanais le sont avec l'article de connexion : *verð* (*o lapsi i verð*, le crayon jaune, *o stilolapsi o/ o i vərði*, le stylo/ jaune, *i nđyra e verð*, la couleur jaune) et *mesme* (*ki patata e mesme*, à la pomme de terre moyenne). Ce n'est jamais le cas d'autres adjectifs tels que *ješili* (*i kapsa i ješili*, la pince à linge verte, *o lapsi o ješili*, le crayon vert), qui font partie de la classe 1 des adjectifs albanais, celle qui ne connaît pas l'article de connexion.

Les adjectifs empruntés en syntagme polydéfini ne sont jamais intégrés à la morphologie héritée. Soit ils suivent le modèle de la langue d'origine, lorsque celui-ci est décelable (article de connexion albanais importé), soit ils se distinguent par l'absence de morphologie (invariabilité en genre).

Tableau 75 : Origine et type sémantique des adjectifs des syntagmes polydéfinis des tests 1 et 2 (témoins exclus)

	Aspect				**Taille**					**Autres**
Adjectif	*čhindo*	*bango*	*uštime*	*pičhindo*	*baro*	*tikno*	*xurdo*	*učo*	*mesme*	*vaver*
Traduction	coupé	tordu	pointu	non-coupé	grand	petit	tout petit	haut	moyen	autre
Origine	héritée	héritée	héritée	héritée	héritée	héritée	héritée	héritée	albanaise	héritée
N occurrences (%)	10 (55,55 %)	6 (33,33 %)	1 (5,56 %)	1 (5,56 %)	108 (57,45 %)	69 (37,70 %)	8 (4,26 %)	2 (1,06 %)	1 (0,53 %)	19 (100 %)
Total	18 (100%)				188 (100%)					19 (100 %)

	Couleur												
Adjectif	*lolo*	*parno*	*kalo*	*roz*	*ješili*	*saria*	*portokali*	*blu*	*žolto*	*vərð*	*gri*	*višnja*	*verð + žolto*[1]
Traduction	rouge	blanc	noir	rose	vert	jaune	orange	bleu	jaune	jaune	gris	rouge cerise	jaune + jaune
Origine	héritée	héritée	héritée	internatio	albanaise	turque	internatio	internatio	slave	albanaise	internatio	slave	albanaise + slave
N occurrences (%)	44 (22,56 %)	34 (17,44 %)	28 (14,36 %)	25 (12,82 %)	20 (10,26 %)	14 (7,18 %)	9 (4,61 %)	6 (3,08 %)	6 (3,08 %)	5 (2,56 %)	2 (1,03 %)	1 (0,51 %)	1 (0,51 %)
Total	195 (100 %)												

[1] Il s'agit d'un cas d'ajout ou d'auto-correction sémantique : *paše ko manykyri i ve/ ə i verð/ žolto*, à côté du vernis à ongle jau/ euh jaune$_{albanais}$, jaune$_{slave}$.

Dans l'ensemble des adjectifs (Tableau 76), on trouve lexicalement à peine plus d'adjectifs hérités (52,17 %) que d'adjectifs empruntés (47,83 %), qui sont constitués quasiment exclusivement d'adjectifs de couleur. En revanche, les adjectifs hérités sont massivement plus employés (78,76 %) que les adjectifs empruntés (21,24 %) qui sont donc seulement d'un usage ponctuel et restreint aux adjectifs de couleur.

Tableau 76 : Synthèse de l'origine des adjectifs en fonction de leur sémantique des syntagmes polydéfinis des tests 1 et 2 (témoins exclus)

	Aspect		Taille		Autres		Couleur		Total	
Origine	**héritée**	**empr.**	**héritée**	**empr.**	**héritée**	**empr.**	**héritée**	**empr.**	**héritée**	**empr.**
N adjectifs (%)	4 (100 %)	0 (%)	4 (80 %)	1 (20 %)	1 (100 %)	0 (0 %)	3 (23,08 %)	10 (76,92 %)	12 (52,17 %)	11 (47,83 %)
Total (%)	4 (100 %)		5 (100 %)		1 (100 %)		13 (100 %)		23 (100 %)	
N occurrences (%)	18 (100 %)	0 (0 %)	187 (99,47 %)	1 (0,53 %)	19 (100 %)	0 (0 %)	106 (53,64 %)	88 (45,36 %)	330 (78,76 %)	89 (21,24 %)
Total (%)	18 (100 %)		188 (100 %)		19 (100 %)		194[1] (100 %)		419 (100 %)	
% des totaux	18 (4,30 %)		188 (44,87 %)		19 (4,53 %)		194 (46,30 %)		419 (100 %)	

On voit donc que l'origine empruntée d'un adjectif n'empêche pas son emploi dans une structure polydéfinie. Il reste cependant très limité, à 21,24 % des occurrences de syntagmes DEF N DEF A. Cela est-il dû à une préférence lexicale, comme nous l'avons suggéré, ou bien à la polydéfinitude elle-même, qui inhiberait leur usage ?

2.4.3 Origine des adjectifs dans tous les syntagmes définis

Pour répondre à cette question, il convient d'observer la syntaxe des adjectifs dans les autres syntagmes définis : c'est l'objet du Tableau 77.

[1] Nous avons ici exclu du compte l'occurrence mixte.

Tableau 77 : Synthèse de l'origine des adjectifs dans tous les SN définis des tests 1 et 2 (témoins exclus)

		Aspect		Taille		Autres		Couleur		Total		Total des totaux	
	Origine	héritée	empr.	héritée	empr.	héritée	empr.	héritée	empr.	héritée	empr.	héritée	empr.
DEF N DEF A	N occurrences	18 (100 %)	0 (0 %)	187 (99,47 %)	1 (0,53 %)	19 (100 %)	0 (0 %)	106 (53,64 %)	88 (45,36 %)	330 (78,76 %)	89 (21,24 %)	827 (76,22 %)	258 (23,78 %)
	Total occurrences (%)	18 (100 %)		188 (100 %)		19 (100 %)		194[1] (100 %)		419 (100 %)			
	% des totaux	18 (4,30 %)		188 (44,87 %)		19 (4,53 %)		194 (46,30 %)		419 (100 %)			
DEF N A	N occurrences	11 (100 %)	0 (0 %)	95 (100 %)	0 (0 %)	16 (100 %)	0 (0 %)	57 (30,48 %)	130 (69,52 %)	179 (57,93 %)	130 (42,07 %)		
	Total occurrences (%)	11 (100 %)		95 (100 %)		16 (100 %)		187[2] (100 %)		309 (100 %)			
	% des totaux	11 (3,56 %)		95 (30,74 %)		16 (5,18 %)		187 (60,52 %)		309 (100 %)			
DEF A N	N occurrences	23 (100 %)	0 (0 %)	194 (100 %)	0 (0 %)	14 (70 %)	6 (30 %)	69 (71,13 %)	28 (28,87 %)	300 (89,82 %)	34 (10,18 %)		
	Total occurrences (%)	23 (100 %)		194 (100 %)		20 (100 %)		97 (100 %)		334 (100 %)			
	% des totaux	23 (6,89 %)		194 (58,08 %)		20 (5,99 %)		97 (29,04 %)		334[3] (100 %)			
complexes	N occurrences	0 (0 %)	0 (0 %)	9 (100 %)	0 (0 %)	3 (100 %)	0 (0 %)	6 (54,55 %)	5 (45,45 %)	18 (78,26 %)	5 (21,74 %)		
	Total occurrences (%)	0 (0 %)		9[4] (100 %)		3 (100 %)		11[5] (100 %)		23 (100 %)			
	% des totaux	0 (0 %)		9 (39,13 %)		3 (13,04 %)		11 (47,83 %)		23[6] (100 %)			
Total	% des totaux des totaux	52 (4,79 %)		486 (44,79 %)		58 (5,35 %)		489 (45,07 %)		1085[7] (100 %)		1085 (100 %)	

[1] Nous avons ici exclu du compte 1 occurrence mixte.
[2] Nous avons ici exclu du compte 6 occurrences mixtes.
[3] Nous avons ici exclu du compte 2 occurrences mixtes dans le type d'adjectif (aspect + taille), toutes deux d'origine héritée.
[4] Nous avons ici exclu du compte 1 occurrence mixte.
[5] Nous avons ici exclu du compte 6 occurrences mixtes.
[6] Nous avons ici exclu du compte 20 occurrences mixtes tant dans le type d'adjectif (5 autres + couleur, 8 autres + taille / taille + autre, 5 couleur + taille / taille + couleur, 2 taille + aspect) que dans l'origine (12 hérités, 8 mixtes).
[7] Nombre auquel il faut donc ajouter les 1+6+2+1+6+20 (= 36) occurrences retirées pour cause de mixité, pour retrouver le total de 1121 occurrences prononcées par les participants albanais dans les deux versions du test.

Le Tableau 77 montre que les trois quarts des adjectifs employés en syntagme défini dans le test sont d'origine indo-aryenne (76,22 %)[1]. Les adjectifs hérités sont très majoritaires lorsque le locuteur emploie la structure canonique DEF A N (89,82 %) et la structure polydéfinie DEF N DEF A (78,76 %), ainsi que dans les structures complexes (78,26 %). En revanche, lorsque le locuteur emploie une structure non-canonique DEF N A, c'est pour employer une majorité d'adjectifs empruntés (57,93 %). Il semble qu'il existe une hiérarchie d'*inheritedness* (littéralement « héritéité », le fait d'être hérité), héritage, où les trois structures définies forment une échelle de la plus « romani » à la plus « liée au contact » (Figure 45).

DEF A N > DEF N DEF A > DEF N A

Figure 45 : Échelle d'héritage des syntagmes définis

Lorsqu'un adjectif est emprunté, il est placé tout au bas de cette échelle, en DEF N A. Cette hypothèse contredit l'assertion de Matras (2009 : 188) selon laquelle les adjectifs empruntés tendent à être intégrés en position adnominale dans la langue réceptrice. Cela rappelle en revanche fortement le phénomène de compartimentation (scission de la grammaire entre une morphologie « héritée » et une « empruntée ») (Matras 2002 : 209-213), qui se situerait cette fois sur le plan de la syntaxe. Pour vérifier cette hypothèse, il convient d'inverser la perspective en observant les structures où sont employés de préférence les adjectifs hérités et empruntés : c'est l'objet du Tableau 78.

Tableau 78 : Préférences structurelles des adjectifs en fonction de leur origine dans les SN définis des tests 1 et 2 (témoins exclus)

	A hérités	A empruntés	Total
DEF N DEF A	330 (39,90 %)	89 (34,49 %)	419 (38,62 %)
DEF N A	179 (21,64 %)	130 (50,39 %)	309 (28,48 %)
DEF A N	300 (36,28 %)	34 (13,18 %)	334 (30,78 %)
complexes	18 (2,18 %)	5 (1,94 %)	23 (2,12 %)
Total	827 (100 %)	258 (100 %)	1085 (100 %)

Le Tableau 78 montre que les adjectifs hérités « préfèrent » DEF N DEF A (39,90 %), suivis de près par DEF A N (36,28 %), contrairement à DEF N A qui présente un taux très bas (21,64 %). Au contraire, les adjectifs empruntés « préfèrent » DEF N A (50,39 %), suivis de loin par DEF N DEF A (34,49 %) ; ils sont à peine employés dans les structures DEF A N (13,18 %). Cela montre qu'il est quasiment contradictoire, en romani, d'importer un adjectif et de l'antéposer, et cela confirme la hiérarchie de la Figure 45 et l'hypothèse de la compartimentation syntaxique. Les pourcentages très hauts de DEF

[1] Nombres et pourcentages verts dans le Tableau 77.

N DEF A sont certainement dus à la tâche, qui a bien fonctionné puisqu'elle visait à faire produire cette structure précise. Celle-ci confirme son statut intermédiaire entre la structure la plus « héritée » (et la plus canonique) et la structure la plus « empruntée ». Elle semble même pencher plutôt du coté hérité, si l'on en croit son pourcentage plus élevé d'emprunt chez les adjectifs hérités que chez les adjectifs empruntés.

S'il est contradictoire, en romani, d'importer un adjectif et de l'antéposer, puisque cela viole la hiérarchie, comment expliquer les trente-quatre occurrences DEF A N où l'adjectif n'est pas d'origine romani ? C'est l'objet du Tableau 79.

Tableau 79 : Analyse des occurrences d'adjectifs empruntés dans les syntagmes DEF A N

Adjectif	Traduction	Type	Occurrences	Origine	Variété	Classe d'âge
ble / bluja	bleu	couleur	3	grec / albanais	arli	40-59 ans + 60 ans
ješil / ješili / ješilka	vert	couleur	5	turc & albanais	arli arli & čergar I	toutes classes
portokali	orange	couleur	1	panbalkanique	arli & čergar I	15-39 ans
ros / roza	rose	couleur	7	international	arli arli & čergar I	15-39 ans 40-59 ans + 60 ans
sarija	jaune	couleur	8	turc	arli arli & čergar I	15-39 ans 40-59 ans + 60 ans
zelino	vert	couleur	4	panslave	arli[1]	40-59 ans
saji	droit	autre	6	turc	arli	40-59 ans

A l'exception des mots de large extension tels que *portokali* et *ros*, ainsi que *jeşil*, qui peut être aussi bien albanais que turc (le premier l'ayant emprunté au second), on constate que tous ces adjectifs proviennent de langues qui par défaut antéposent l'adjectif : grec, turc, slave. Ces quelques cas montrent qu'un schéma de convergence syntaxique (adaptation des ressources héritées à un modèle extérieur) existe également pour les adjectifs hérités – même s'il est limité à un petit nombre d'occurrences. Notons que tous ces adjectifs sont employés par des locuteurs parlant l'arli ou deux variétés dont l'arli. La classe d'âge ne nous éclaire pas beaucoup – si ce n'est que les plus jeunes, moins de quinze ans, sont presque absents du tableau. Nous reviendrons plus loin sur l'importance de ces variables sociales.

[1] L'usage de cet adjectif n'est pas habituel dans cette variété, qui exprime « vert » plutôt avec *ješil*. Le locuteur en question a acquis ce mot à l'occasion du test : se refusant à utiliser *ješil*, dont l'origine albanaise potentielle est transparente, il a stoppé le test, est allé chercher un dictionnaire du romani (le seul disponible en albanais, Courthiade 2004). Il y a trouvé « *zèleno* », qu'il a transformé en *zelino*, plaçant l'accent soit sur la première soit sur la dernière syllabe en fonction des occurrences. Ce locuteur ne parle pas de langues slaves. Le fait qu'il l'ait trouvé dans un ouvrage de référence peut porter à croire que le terme fait partie du romani fondamental et n'est donc pas emprunté, d'où la possibilité de l'antéposer.

Pour vérifier s'il y a compartimentation ou convergence, il nous faut étudier les adjectifs postposés empruntés à des langues antéposant l'adjectif. Les adjectifs empruntés et employés en DEF N A sont analysés dans le Tableau 81.

Les adjectifs sont ici bien plus nombreux – en revanche il s'agit exclusivement d'adjectifs de couleur. Ils sont empruntés à des langues très diverses, y compris à des langues qui postposent l'adjectif (albanais, roman). Les adjectifs postposés empruntés à des langues qui antéposent l'adjectif, ce qui va dans le sens de la compartimentation plutôt que de la fusion, sont les suivants : *ble* (grec, une occurrence), *sarija* (turc, six occurrences), *zeleno* (slave, quatorze occurrences), *žolto* (slave, dix occurrences), *zlato* (slave, une occurrence). Ces occurrences ne sont pas nombreuses mais font le pendant des adjectifs empruntés en DEF A N. Notons qu'ici, les adjectifs sont prononcés par des personnes parlant toutes les variétés, et que la représentation des locuteurs de plus de soixante ans est légèrement moindre que celle des autres classes d'âge.

Sous quelle forme se présentent les adjectifs empruntés ? Le Tableau 74 présentait ceux des syntagmes polydéfinis, voyons maintenant ceux des syntagmes monodéfinis et complexes.

Tableau 80 : Morphologie des adjectifs empruntés dans les SN monodéfinis et complexes des tests 1 et 2 (témoins exclus)

Formes	Traduction	Morphologie
ble	bleu	M.SG
bluja		
blu		M.SG, F.SG
gri	gris	M.SG
ješil	vert	M.SG, F.SG
ješili		M.SG, F.SG
ješilka		F.SG
kafe	café	M.SG
portokali	orange	M.SG, F.SG
portokalija		M.SG
ros	rose	M.SG
roz		M.SG, F.SG
roza		
sarija	jaune	M.SG

Formes	Traduction	Morphologie
verð	jaune	M.SG
vərð		
verða		F.SG
zeleno	vert	M.SG, F.SG
zeleni		F.SG
zelino		M.SG, F.SG
želino		M.SG
želini		F.SG
zlato	doré	F.SG
žolto	jaune	M.SG, PL
zurlo	jaune	M.SG
zurno		
fundi	dernier	M.SG
mesmja	moyen	F.SG

Tableau 81 : Analyse des occurrences d'adjectifs empruntés dans les syntagmes DEF N A

Adjectif	Traduction	Type	Occurrences	Origine	Variétés	Classes d'âge
ble / blu / bluja	bleu	couleur	18	grecque / albanaise	arli, arli & čergar I, néo-locuteur arli	-15, 15-39, + 60 ans
gri	gris	couleur	3	internationale	mečkar	15-39 ans
ješil / ješili / šili	vert	couleur	21	turque & albanaise	toutes sauf çergar, mečkar & čergar I, mečkar & bamlija	toutes classes
kafe	café	couleur	1	albanaise	arli	40-59 ans
portokali / portokalija	orange	couleur	25	panbalkanique	toutes sauf mečkar & čergar I et mečkar & bamlija	toutes classes
ros / roz / roza	rose	couleur	26	internationale	toutes sauf çergar, mečkar & čergar I, mečkar & bamlija	toutes classes
sarija	jaune	couleur	6	turque	arli & čergar I, čergar I	-15 ans, 15-39 ans
verð	jaune	couleur	3	albanaise	arli & čergar I, mečkar	-15 ans, 15-39 ans
zeleno/zeleni/zelino /želino/žilino/ žilini	vert	couleur	14	panslave	arli[1], arli & čergar I, meçkar, čergar I, mečkar & bamlija	40-59 ans
žolto	jaune	couleur	10	slave du sud	meçkar, mečkar & bamlija	15-39 ans, 40-59 ans
zlato	doré	couleur	1	panslave	mečkar & čergar I	40-59 ans
zurlo / zurno	jaune	couleur	2	romane (?)	mečkar & bamlija	15-39 ans

[1] Cf. note 1, p. 337.

Le corpus mononéfini permet cette fois de tester la différence singulier ~ pluriel : *o lapsja žolto*, les crayons jaunes, qui montre que *žolto* ne varie pas non plus en nombre (s'il avait été intégré à la morphologie romani, on aurait **žolte*). Le corpus permet aussi de tester la différence cas direct ~ cas oblique : *ti saji vastesar*, avec la main droite, *te saji thaneste/ musjate*, du côté/de la main droite, qui montre que *saji* ne varie pas non plus en cas (s'il avait été intégré à la morphologie romani, on aurait **saje*).

L'invariabilité en genre de certains adjectifs du corpus polydéfini est confirmée, pour lesquels on n'avait qu'une seule configuration : *blu, bluja, ješil, portokalija*. D'autres formes se voient confirmées comme invariables : *ješili, roz, roza, zeleno.* Difficile de savoir s'il pourrait en être autrement pour *blu, bluja, ješil, ješili, portokali, ros, roz*, qui ne varieraient pas en genre en albanais, au nominatif singulier déterminé. De nouveaux adjectifs sont apparus, qui ne sont employés qu'une fois, ce qui ne permet pas d'établir une comparaison : *kafe, fundi. Saria* reste confiné au masculin, mais on peut savoir qu'il est en réalité invariable, en observant le corpus des syntagmes DEF A, où il est employé pour modifier un nom féminin sous-entendu : *i saria*, où *kapsa*, pince à linge, a été prononcé au syntagme précédent[1].

Notons deux cas de variation morphologique. Le premier concerne l'adjectif *zeleno*, vert. Cet adjectif a été emprunté sous sa forme neutre en slave : le neutre s'y caractérise par la désinence *-o*, homophone de la désinence masculine *-o* du romani. Une réanalyse est donc possible, *zeleno* comme masculin singulier romani plutôt que comme forme empruntée invariable. Cela amène deux locutrices (meçkar) à en déduire une forme *zeleni* où l'adjectif porte la désinence féminine *-i* : l'une prononce *i kapsa e gadengi zeleni*, la pince des vêtements verte[2], l'autre commence par employer l'adjectif sous sa forme invariable *-o*, avant de s'auto-corriger et d'employer la désinence *-i* : *i kàpsea žilino/ žilini/ žilini*, la pince à linge verte$_M$/ verte$_F$/ verte$_F$.

Le second cas de variation morphologique concerne l'adjectif *ješil*, vert. Un morphème du féminin singulier autre que *-i* se détache : *-ka*. Emprunté au slave, il permet de féminiser un nom masculin, par exemple *mečkar*, homme meçkar ~ *mečkarka*, femme meçkar, parfois sans qu'il y ait changement de sens, ainsi *kompir = kompirka*, pomme de terre, *topi = topka*, balle/boule. Il est aussi employé comme suffixe féminin de noms par ailleurs féminins, par exemple *domatka*, doublet de *domata*, tomate, *patatka = patata*, pomme de terre, *kapska = kapsa*, pince à linge[3]. On le trouve aussi en fin de mots féminins qui n'ont pas d'équivalent masculin (*kultulka*, pomme de terre). Une locutrice, voulant féminiser l'adjectif *ješil*, n'a pas pu utiliser la morphologie romani héritée, car l'ajout de *-i* n'est pas significatif : *ješili* est une variante phonétique, ambigène, du même adjectif[4]. Elle a donc

[1] « *i saria* » est dit en réponse à une demande de précision de la part du participant B (*solicited new-information focus*).

[2] Ce n'est pas un hapax : cette locutrice prononce aussi un syntagme indéfini *jek kapsa/ zeleni*, une pince à linge/ verte.

[3] Notons que dans ces cas, le suffixe n'a pas de valeur diminutive, comme il peut en avoir dans les langues slaves. On trouve ainsi *i patatka i bari*, la grosse pomme de terre.

[4] Le *-i* final vise à conserver la compartimentation phonologique du romani : mot à accent final = hérité du stock lexical panromani ≠ mot à accent non-final = mot emprunté aux langues de contact. En effet, l'accent tonique doit toujours rester sur la syllabe d'origine lorsqu'un lexème est importé (romani [ˈʒol.to] et non [ʒol.ˈto], du macédonien & bulgare neutre singulier [ˈʒol.to]). Cela pose un problème lorsque la position d'origine en question est la syllabe finale de mot, à l'instar des mots romanis hérités. Ainsi le turc [sa.ˈɾɯ] n'a-t-il pas donné en romani [sa.ˈɾɯ] ni [sa.ˈɾi], mais [sa.ˈɾi.ja] après ajout

cherché un suffixe qui joue le rôle d'adaptateur (*loan-adaptation marker*, Matras 2002), d'où *i ješilka kapse*, la pince à linge verte. On peut penser que c'est le même raisonnement qui amenait une autre locutrice à prononcer deux occurrences de *i kapsa i ješilka*, la pince à linge verte. Notons qu'une troisième locutrice a prononcé le syntagme avec ellipse du nom *anglal ti ješilka*, derrière la verte. Notons que toutes trois parlent la variété arli.

On retrouve en syntagmes monodéfinis et complexes des adjectifs empruntés à l'albanais avec l'article de connexion : *fundi* (*o kalemi o xurdo i fundi*, le tout petit dernier crayon), et *mesmja* (*i patàta i/ i tikini e mesmja*, la pomme de terre/ petite moyenne). Notons que *verð* se rencontre trois fois sans article de connexion : *o lapsi verð*, le crayon jaune. Dans l'occurrence *i bojàva i verða*, la couleur jaune, le deuxième article *i* peut s'interpréter là aussi comme un article romani dupliqué, car l'article de connexion albanais attendu serait *e* au féminin – le *-a* final de l'adjectif pourrait éventuellement être interprété comme une marque féminine.

Les adjectifs empruntés en syntagme monodéfini ne sont donc quasiment jamais intégrés à la morphologie héritée. Soit ils suivent le modèle de la langue d'origine, lorsque celui-ci est décelable (article de connexion albanais importé), soit ils se distinguent par l'absence de morphologie (invariabilité en genre, nombre, cas, éventuellement l'absence d'article de connexion albanais). On a remarqué cependant deux tentatives d'intégrer les adjectifs à la morphologie du romani : grâce à la désinence adjectivale héritée (*zeleni*, verte) ou grâce à une désinence nominale *-ka* (*ješilka*, verte).

2.4.4 Sémantique des adjectifs dans tous les syntagmes définis

Les adjectifs de notre corpus sont tous restrictifs, car le test de manipulation a été conçu à dessein. Est-il possible de déterminer si l'usage de la polydéfinitude est favorisé par la sémantique de l'adjectif (aspect, taille, couleur….), ou encore par son origine (héritée ou empruntée) ? Le Tableau 77 montre que les adjectifs les plus employés en syntagme défini sont ceux de couleur et de taille : ils représentent respectivement 45,07 % et 44,79 % des occurrences. On trouve des taux similaires en syntagme polydéfini : respectivement 46,30 % et 44,87 % des occurrences (Tableau 76). Cela est dû à la nature de la tâche, qui limitait évidemment le type d'adjectif en jeu, mais aussi aux objets que j'avais choisis de faire manipuler ainsi qu'aux types de propriétés que j'ai choisi de faire contraster.

Les Tableau 127 et Tableau 128 présentent l'ensemble les adjectifs employés dans le corpus adjectifs utilisés au moins une fois pour chaque structure. On peut constater que les adjectifs qui sont employés dans la plus grande variété de structures sont *baro*, grand, *ješil*, vert, *kalo*, noir, *lolo*, rouge, *parno*, blanc, *roz*, rose, *tikno*, petit, *vaver*, autre, *zeleno*, vert, *žolto*, jaune. On voit qu'il y a tous les types d'adjectif sauf les adjectifs d'aspect.

---

de la syllabe [ja] finale. Les deux contraintes sont donc satisfaites : respect de la place de l'accent tonique issu de la langue d'emprunt, mais accent non-final une fois le mot emprunté. Cela explique pourquoi on trouve *ješili* et *ješilka* mais peu de *ješil* : l'emprunt est fait au turc/albanais *yeşil/jeshil*.

Les différents types d'adjectif n'apparaissent cependant pas de manière uniforme selon les structures : les adjectifs de couleur apparaissent de manière très majoritaire dans la structure DEF N A (60,52 %)[1], moins dans la structure polydéfinie (46,30 %), et moins encore dans la structure DEF A N (29,04 %). Les adjectifs de taille présentent des proportions inversement similaires : 58,08 % dans la structure DEF A N, 44,87 % dans la structure DEF N DEF A et seulement 30,74 % dans la structure DEF N A. Y a-t-il un lien entre la sémantique inhérente des adjectifs et la fonction des différentes structures, ou l'explication par l'origine de l'adjectif suffit-elle ?

Pour vérifier cette hypothèse, il convient d'inverser la perspective en observant les structures où sont employés de préférence les différents types d'adjectifs : c'est l'objet du Tableau 82.

Tableau 82 : Préférences syntaxiques des adjectifs en fonction de leur sémantique dans les SN définis des tests 1 et 2 (témoins exclus)

	Aspect	Taille	Autres	Couleur	Total
DEF N DEF A	18 (34,62 %)	188 (38,68 %)	19 (32,76 %)	194 (39,67 %)	419 (38,62 %)
DEF N A	11 (21,15 %)	95 (19,55 %)	16 (27,59 %)	187 (38,24 %)	309 (28,48 %)
DEF A N	23 (44,23 %)	194 (39,92 %)	20 (34,48 %)	97 (19,84 %)	334 (30,78 %)
complexes	0 (0 %)	9 (1,85 %)	3 (5,17 %)	11 (2,25 %)	23 (2,12 %)
Total	52 (100 %)	486 (100 %)	58 (100 %)	489 (100 %)	1085 (100 %)

Le Tableau 82 montre deux types opposés de syntaxe. Dans le premier, DEF A N représente la structure préférée, suivie de DEF N DEF A et de DEF N A. C'est le cas pour les adjectifs d'aspect, de taille et d'un autre sens. Parmi eux, l'écart le plus important se trouve chez les adjectifs d'aspect : 21,15 % (DEF N A) à 44,23 % (DEF A N). Dans le deuxième type de syntaxe, DEF N DEF A représente la structure préférée (39,67 %), suivie de très près par DEF N A (38,24 %), et de très loin par DEF A N (19,84 %) : il s'agit des adjectifs de couleur.

Les adjectifs de taille et « autres » présentent cependant une situation remarquable. Ils suivent le patron des adjectifs d'aspect pour la hiérarchie DEF A N > DEF N A, mais la différence est minime entre DEF A N (respectivement 39,92 % taille, 34,48 % autres) et DEF N DEF A (respectivement 38,68 % taille, 32,76 % autres). On peut dire qu'ils suivent le type de syntaxe des adjectifs d'aspect et respectent la hiérarchie d'héritage, mais avec un problème : le taux de DEF N DEF A est « trop » élevé, la hiérarchie DEF A N > DEF N DEF A n'est « pas assez » clairement marquée. En outre, les adjectifs autres présentent un taux de DEF N A très important (27,59 % alors qu'il est autour de 20 % chez ceux d'aspect et de taille), ce qui le rapproche des adjectifs de couleur (38,24 %). Les adjectifs de taille et autres représentent en fait un troisième type de syntaxe, intermédiaire. On peut résumer le Tableau 82 et le

[1] Nombres et pourcentages bleus dans le Tableau 77, p. 339.

continuum de types syntaxiques que vous venons de décrire en classant les préférences syntaxiques des différents types d'adjectifs en Figure 46 :

aspect : DEF A N > DEF N DEF A > DEF N A

taille & autres : DEF A N ≥ DEF N DEF A > DEF N A

couleur : DEF N DEF A ≥ DEF N A > DEF A N

Figure 46 : Trois types de préférences syntaxiques en fonction du type d'adjectif

Dans quelle mesure ces trois types de syntaxe se laissent-ils expliquer par l'origine des adjectifs, et dans quelle mesure faut-il chercher une autre explication ? Il faut pour cela rechercher si ces types d'adjectifs respectent la hiérarchie d'héritage de la Figure 45.

Tableau 83 : Synthèse de l'origine des adjectifs en fonction de leur sémantique dans les SN définis des tests 1 et 2 (témoins exclus)

	Aspect		**Taille**		**Autres**		**Couleur**		**Total**	
Origine	**héritée**	**empr.**	**héritée**	**empr.**	**héritée**	**empr.**	**héritée**	**empr.**	**héritée**	**empr.**
N occurrences (%)	52 (100 %)	0 (0 %)	485 (99,79 %)	1 (0,21 %)	52 (89,66 %)	6 (10,34 %)	238 (48,67 %)	251 (51,33 %)	**827 (76,22 %)**	**258 (23,78 %)**
Total (%)	52 (100 %)		486 (100 %)		58 (100 %)		489 (100 %)		1085 (100 %)	
% des totaux	**52 (4,79 %)**		**486 (44,79 %)**		**58 (5,35 %)**		**489 (45,07 %)**		**1085 (100 %)**	

Le Tableau 83 nous permet de classer les types d'adjectifs en fonction du plus ou moins grand nombre d'occurrences héritées ou empruntées (Figure 47) :

aspect > taille > autres > couleur

Figure 47 : Proportion décroissante d'adjectifs hérités en fonction des types sémantiques

Il montre aussi que les adjectifs d'aspect présentent un taux de 0 % d'occurrences d'adjectifs empruntés, tandis que les adjectifs de couleur présentent un taux de 51,33 %. Cela peut suffire à expliquer les deux types de syntaxe différents qu'ils présentent : tous deux respectent la hiérarchie d'héritage et tous deux se situent à un pôle opposé de l'échelle.

Les adjectifs de taille posent en revanche un problème : ils sont pour ainsi dire quasiment toujours hérités (1 occurrence d'emprunt, 0,21 %), à l'instar des adjectifs d'aspect… et pourtant leur type syntaxique diffère de celui des adjectifs puisqu'ils présentent un taux de DEF N DEF A (38,68 %) quasiment égal au taux de DEF A N (39,92 %) (Tableau 82). Ils présentent un type syntaxique

partiellement similaire à celui des adjectifs de couleur alors qu'ils sont hérités. Ils violent l'échelle d'héritage à la charnière entre DEF A N et DEF N DEF A.

Les adjectifs autres présentent un petit taux d'adjectifs empruntés (10,34 %) qui pourrait expliquer pourquoi ils occupent une position intermédiaire entre le type syntaxique « adjectif d'aspect » et le type syntaxique « adjectif de couleur ». Cela expliquerait pourquoi ils se situent vers le pôle « emprunté » de la hiérarchie d'héritage, avec un taux de DEF N DEF A et de DEF N A tous deux importants. Cependant cette explication ne tient pas, car ces 6 occurrences d'adjectifs empruntés sont toutes en syntagmes DEF A N avec l'adjectif d'origine turque *saji*, droit, (Tableau 79), qui contribuent au taux élevé de DEF A N (34,48 %) et non de DEF N DEF A (32,76 %) et de DEF N A (27,59 %). Sans cela, l'adjectif *vaver*, autre, pose en fait le même problème que les adjectifs de taille : il est d'origine héritée mais présente lui aussi un taux très haut de DEF N DEF A (36,54 % si l'on fait le calcul en ne tenant compte que des 52 occurrences de *vaver*). Il présente aussi un taux très haut de syntagme DEF N A (30,77 % *vaver*), plus haut que le taux de DEF A N (26,92 % *vaver*). Il présente un type syntaxique similaire à celui des adjectifs de couleur alors qu'il n'est pas hérité. Il viole quasiment l'échelle d'héritage à la charnière entre DEF A N et DEF N DEF A si l'on tient compte des 6 occurrences de *saji*, et la viole même complètement si l'on omet ces occurrences, présentant alors une préférence DEF N A > DEF N DEF A > DEF A N. On peut résumer le Tableau 82 et le Tableau 83 en mettant en évidence ce qui, dans la Figure 46, pose problème :

Aspect : DEF A N > DEF N DEF A > DEF N A
Taille & autres : DEF A N ≥ DEF N DEF A > DEF N A
Vaver seul : DEF N DEF A > DEF N A > **DEF A N**
Couleur : DEF N DEF A ≥ DEF N A > **DEF A N**

Figure 46 : Préférences syntaxiques des différents types d'adjectifs

Nous avons vu que les « problèmes » posé par les adjectifs de couleur n'en sont pas : ils s'expliquent simplement par la part importante d'adjectifs empruntés (Figure 47), qui les situe à l'opposé de l'échelle d'héritage (Figure 45). En revanche, les adjectifs de taille et autres, notamment *vaver*, posent un problème qui ne se laisse pas expliquer par le contact de langues. Cela requiert d'aller au-delà de la seule « sémantique » des adjectifs.

2.4.5 La combinaison sémantique adjectif – nom

Nous avons distingué pour l'instant quatre types d'adjectifs en fonction de leur sens : aspect, taille, autres, couleur. Cette distinction ne s'est pas révélée suffisante en ce qui concerne les adjectifs de taille, et l'origine étymologique n'a pas pu nous aider non plus. Quelle est donc la particularité des adjectifs de taille ?

Il convient pour cela de se pencher sur l'interprétation d'un syntagme nominal avec adjectif de taille versus avec un adjectif d'aspect, puisque l'interprétation d'un syntagme résulte de la lecture combinée de l'adjectif et du nom. Nous nous appuyons pour cela sur la synthèse des travaux en

sémantique formelle établie par Morzycki (2015) ainsi que ceux de Stavrou (1999) et Cinque (2010 : 20) (cf. section 2.3, p.82).

Les adjectifs de **couleur** sont intersectifs : *o kalemi o lolo*, le crayon rouge, implique que le référent appartient à l'ensemble des crayons et à l'ensemble des entités rouges. C'est pourquoi on dit qu'il se trouve à l'intersection des deux ensembles, donc que le syntagme a une lecture intersective. Ces adjectifs sont *object-oriented* au sens de Stavrou (1999). Les adjectifs d'**aspect** sont eux aussi intersectifs et *object-oriented* : *i domatka i uštime*, la tomate pointue, implique que le référent se trouve à l'intersection des tomates et des entités pointues. On a vu en section 2.4, p. 329, que les adjectifs de **taille**, quant à eux, semblent à lecture subsective, mais qu'ils sont en réalité à lecture intersective. En effet, *i kultulka i tikni*, la petite pomme de terre, peut être plus grosse que *i sir i bari*, la grosse tête d'ail. Le sens de *bari*, grande, est à interpréter comme « grande-pour-une-tête-d'ail », en relation étroite avec le sens du nom car la dimension d'un élément est relative. Les adjectifs que nous avons classés dans la catégorie **autres** font eux aussi débat – à ceci près que leur cas n'a pas toujours été débattu par la littérature, notamment celui de l'adjectif *autre*. Dans le syntagme *i patata i vaver*, l'autre pomme de terre, le référent appartient à l'ensemble des {pommes-de-terre}, mais existe-t-il un ensemble des {autres entités} ? On a considéré l'interprétation du syntagme nominal comme intersective en apparence subsective. En effet l'adjectif « autre » implique par définition la présence d'*autres* référents potentiels en contexte. *Vaver*, autre, n'est pas le seul adjectif de la catégorie « autre » : nous y avons classé également six occurrences de *saji*, droit, toutes utilisées en syntagme DEF A N. Cet adjectif (arli *ti saji musi*, à droite, litt. au bras droit), de même que *majt(i)*, gauche (corpus spontané mečkar *ko piko tiro/ i majt*, à ta gauche, litt. à ton épaule/ gauche), est également intersectif en apparence subsectif. Ces adjectifs sont restrictifs car lorsqu'on parle d'un objet situé sur la gauche, c'est nécessairement en relation avec un autre qui est sur la droite. La localisation spatiale présente les mêmes caractéristiques sémantiques que « autre » : le sens est purement contrastif, de manière binaire (*gauche ~ droit*), et l'adjectif représente un sous-ensemble du nom dicté par le contexte d'énonciation. Ces adjectifs sont *speaker-oriented*.

Le Tableau 84 donne un aperçu des interprétations sémantiques des différents adjectifs.

Tableau 84 : Type d'interprétation des adjectifs des syntagmes polydéfinis des tests 1 et 2 (témoins exclus)

		Aspect	Taille	Autres	Couleur
(Morzycki 2015)	Lecture intersective	✓			✓
	Lecture subsective				
	Lecture intersective en apparence subsective		✓	✓	
	Lecture ni intersective ni subsective				
	Lecture privative				
(Stavrou 1999)	Orienté vers l'objet	✓	✓		✓
	Orienté vers le locuteur			✓	

Nous voyons que la polydéfinitude est employée, dans le corpus du test, avec des adjectifs à lecture intersective et intersective en apparence subsective, avec des adjectifs orientés vers l'objet et orientés vers le locuteur. À ce stade, nous ignorons encore si la structure est possible avec des adjectifs à lecture réellement subsective (qui appartiennent à la catégorie *speaker-oriented*), à lecture ni intersective ni subsective (*speaker-oriented*) et à lecture privative (*speaker-oriented*)

On s'aperçoit que les adjectifs de taille et autres se distinguent des adjectifs d'aspect et de couleur : ils sont intersectifs en apparence subsectifs, c'est-à-dire que leur interprétation est plus dépendante de celle du nom ou bien du contexte, que les adjectifs « simplement » intersectifs. Cela pourrait expliquer pourquoi ils présentent un taux de DEF N DEF A plus élevé que celui attendu (notamment pour des adjectifs hérités, Figure 45). La tâche des tests vise à contraster des objets similaires se distinguant par une ou deux propriétés au maximum. Lorsqu'un objet ne se distingue d'un autre que par une propriété peu intrinsèque (la taille, « autre » ou la position droite ~ gauche), la structure polydéfinie, dont la raison d'être initiale est le contraste, peut sembler plus appropriée pour l'identification. Les adjectifs de taille et autres [à lecture intersective en apparence subsective] pousseraient donc plus à l'usage de la polydéfinitude que les adjectifs d'aspect et de couleur [à lecture intersective]. Les adjectifs d'aspect et de couleur [à lecture intersective], en revanche, dont la propriété apparaît comme inhérente au référent, nécessitent moins l'usage de la polydéfinitude. En effet, il s'agit toujours, pour le locuteur, de sélectionner un référent parmi un ensemble de référents similaires, mais au moyen d'une propriété dont la sémantique est moins dépendante de celle du nom.

On constate que cette distinction entre adjectifs est la même que pour l'usage des degrés comparatif et superlatif. Les adjectifs de taille et de localisation spatiale changent plus volontiers de degré (*plus petit, moins grand, le plus à gauche, Alexandre le Grand*) que les adjectifs d'aspect et de couleur (*moins coupé, le plus rouge*). L'usage de la polydéfinitude est donc différencié en fonction du type d'adjectif (Figure 48).

Lecture intersective en apparence subsective > lecture intersective

Figure 48 : Hiérarchie sémantique portant à l'usage de la polydéfinitude

2.4.6 Témoins

On trouve chez les témoins une seule occurrence de DEF N DEF A et 4 occurrences de DEF N A. Les témoins ont donc une syntaxe quasiment *catégorique* de DEF A N. Le Tableau 86 présente les adjectifs prononcés par les témoins dans les syntagmes mono- et polydéfinis, classés en fonction de leur origine, de leur type sémantique et de la structure dans laquelle ils apparaissent.

Près des trois quarts des adjectifs employés en syntagme défini dans le test sont d'origine indo-aryenne (70,37 %)[1], comme les locuteurs albanais. Les adjectifs hérités sont très majoritaires lorsque le locuteur emploie la structure canonique DEF A N (75 %). Les adjectifs hérités se trouvent en revanche en proportion bien moindre pour la structure DEF N A (40 %). L'échelle d'héritage est donc

[1] Nombres et pourcentages verts dans le Tableau 86, p. 352.

valable pour les témoins, dans des proportions encore plus tranchées que dans les variétés albanaises : à la structure canonique DEF A N correspond une majorité d'adjectifs hérités, à la structure non-canonique DEF N A correspond une majorité d'adjectifs empruntés. Il convient cependant de rester prudent, car les occurrences, notamment de DEF N A, sont en très petit nombre.

On ne trouve chez les témoins qu'une seule occurrence de syntagme polydéfini, *o lapsi/ to portokal*, le crayon/ orange. L'adjectif de couleur est emprunté, certainement ici au macédonien портокал, *portokal*, qui est l'autre langue principale du locuteur – mais il existe aussi sous cette forme dans les autres langues balkaniques, par exemple albanais *portokall*, grec πορτοκάλι, *portokàli* et bulgare портокал, portokal).

De quel type sont les adjectifs empruntés ? On en trouve deux de taille, un 'autre' et treize de couleur : la situation est ici la même que chez les locuteurs albanais, avec une majorité d'adjectifs de couleur.

Dans quelle structure sont employés les adjectifs empruntés ? On en trouve un en DEF N DEF A, trois en DEF N A, et douze en DEF A N. Cela viole l'échelle d'héritage, puisque cela signifie qu'un adjectif emprunté est quasiment toujours intégré dans la syntaxe héritée. Il convient de regarder plus précisément quels sont ces douze adjectifs empruntés antéposés : c'est l'objet du Tableau 85.

Tableau 85 : Analyse des occurrences d'adjectifs empruntés antéposés chez les témoins

Adjectif	Traduction	Type	Occurrences	Origine	Variété	Classe d'âge
plavo	bleu	couleur	4	slave du sud	kovački	15-39 ans
pomarandža	orange	couleur	1	slave	džambazo	15-39 ans
zeleni	vert	couleur	4	panslave	xoraxane	40-59 ans
žolto	jaune	couleur	1	slave du sud	džambazo	15-39 ans
srednjo	moyen	taille	2	slave du sud	xoraxane	40-59 ans

On constate que la totalité de ces adjectifs est d'origine slave, langues qui antéposent l'adjectif. Il s'agit donc en fait d'une fusion syntaxique, dont participent les trois témoins. Pour vérifier s'il y a compartimentation ou fusion, il nous faut étudier les quatre adjectifs empruntés postposés (DEF N DEF A et DEF N A), qui font l'objet du Tableau 87.

Tableau 86 : Synthèse de l'origine des adjectifs dans tous les SN définis des témoins

		Aspect		Taille		Autres		Couleur		Total		Total des totaux	
	Origine	héritée	empr.	héritée	empr.	héritée	empr.	héritée	empr.	héritée	empr.	héritée	empr.
DEF N DEF A	N occurrences	0 (0 %)	0 (0 %)	0 (0 %)	0 (0 %)	0 (0 %)	0 (0 %)	0 (0 %)	1 (100 %)	0 (0 %)	1 (100 %)		
	Total occurrences (%)	0 (0 %)		0 (0 %)		0 (0 %)		1 (100 %)		1 (100 %)			
	% des totaux	0 (0 %)		0 (0 %)		0 (0 %)		1 (100 %)		1 (100 %)			
DEF N A	N occurrences	0 (0 %)	0 (0 %)	1 (100 %)	0 (0 %)	0 (0 %)	1 (100 %)	1 (33,33 %)	2 (66,67 %)	2 (40 %)	3 (60 %)		
	Total occurrences (%)	0 (0 %)		1 (100 %)		1 (100 %)		3 (100 %)		5 (100 %)		38 (70,37 %)	16 (29,63 %)
	% des totaux	0 (0 %)		1 (20 %)		1 (20 %)		3 (60 %)		5 (100 %)			
DEF A N	N occurrences	3 (100 %)	0 (0 %)	26 (92,86 %)	2 (7,14 %)	5 (100 %)	0 (0 %)	2 (16,67 %)	10 (83,33 %)	36 (75 %)	12 (25 %)		
	Total occurrences (%)	3 (100 %)		28 (100 %)		5 (100 %)		12 (100 %)		48 (100 %)			
	% des totaux	3 (6,25 %)		28 (58,33 %)		5 (10,42 %)		12 (25 %)		48[1] (100 %)			
Total[2]	% des totaux des totaux	3 (5,56 %)		29 (53,70 %)		6 (11,11 %)		16 (29,63 %)		54[3] (100 %)		54 (100 %)	

[1] Nous avons ici exclu du compte 2 occurrences mixtes dans le type d'adjectif (couleur + taille, autres + taille), dont 1 est d'origine héritée, 1 d'origine mixte.

[2] Nous avons ici exclu du compte les 3 syntagmes complexes, qui sont mixtes dans le type d'adjectif (aspect + couleur, aspect + couleur, taille + couleur), tous d'origine héritée.

[3] Nombre auquel il faut donc ajouter les 2 + 3 (= 5) occurrences retirées pour cause de mixité, pour retrouver le total de 59 occurrences prononcées par les témoins.

Tableau 87 : Analyse des occurrences d'adjectifs empruntés postposés chez les témoins

Adjectif	Traduction	Type	Occurrences	Origine	Variété	Classe d'âge
plavo	bleu	couleur	1	slave du sud	džambazo	15-39 ans
roze	rose	couleur	1	international	džambazo	15-39 ans
portokal	orange	couleur	1	panbalkanique	džambazo	15-39 ans
levo	gauche	autres	1	slave du sud	kovački	15-39 ans

On constate ici que l'origine des adjectifs est plus diverse. Ils sont presque tous produits par une même personne, tandis que le locuteur xoraxane ne postpose aucun adjectif[1]. Pour deux adjectifs postposés alors qu'ils sont d'origine slave (et devraient donc être antéposés), on retrouve le modèle de compartimentation syntaxique.

Sous quelle forme se présentent les adjectifs empruntés ? Le Tableau 88 présente les adjectifs de tous les syntagmes poly- et monodéfinis.

Tableau 88 : Morphologie des adjectifs empruntés dans les SN définis produits par les témoins

Traduction	Formes de l'adjectif	Morphologie
orange	*pomarandža*	F.SG
orange	*portokal*	M.SG
rose	*roze*	F.SG
bleu	*plavo*	M.SG
vert	*zeleni*	F.SG
jaune	*žolto*	F.SG
moyen	*srednjo*	M.SG
gauche	*levo*	M.SG.OBL

Le corpus des témoins ne permet pas de tester la variation, puisque tous les adjectifs ne sont employés que dans une configuration. On remarque cependant que *žolto* est employé au féminin, ce qui laisse penser qu'il est invariable, et qu'il y a donc ici compartimentation, comme dans le corpus albanais. On remarque également la forme *zeleni* au féminin, employée à 4 reprises par le locuteur de la variété xoraxane : il semble effectuer la même fusion morphologique que les locutrices albanaises.

2.4.7 Synthèse

Les adjectifs présents dans notre corpus polydéfini sont de poids et de morphologie différents. Ils sont presque tous prédicatifs.

[1] Il produit en revanche trois syntagmes complexes DEF A N DEF A.

Ils sont également d'origine étymologique différente. Il existe une échelle (ou *mapping*) d'héritage (*inheritedness*) des syntagmes définis sur laquelle les locuteurs placent les adjectifs en fonction de leur origine. Plus un adjectif est ressenti comme emprunté, plus il a de probabilité d'être employé en DEF N A plutôt qu'en DEF A N – la polydéfinitude DEF N DEF A se situant à mi-chemin entre les deux. Cela explique pourquoi les adjectifs de couleur sont en très grande majorité postposés (DEF A N ne représente que 19,84 % des occurrences) alors que les adjectifs d'aspect sont plutôt antéposés (DEF A N représente 44,23 % des occurrences) (Tableau 82) : les premiers sont en majorité empruntés (51,33 %), tandis que les seconds sont tous hérités (100 %) (Tableau 83).

L'échelle d'héritage est quasiment enfreinte par les adjectifs de taille et autres, qui présentent un taux de polydéfinitude quasiment égal à DEF A N, alors qu'ils sont presque tous hérités (respectivement 99,79 % et 89,66 %, cf. Tableau 83). Elle est également enfreinte par l'adjectif hérité *vaver*, autre, qui présente des taux anormalement élevés de DEF N DEF A et de DEF N A, mais un taux anormalement bas de DEF A N, comme s'il était un adjectif emprunté ou un adjectif de couleur. C'est que les adjectifs de taille et *autre* appartiennent à une catégorie différente de ceux d'aspect et de couleur sur le plan sémantique : comme nous l'avons montré plus haut, ils sont à lecture intersective en apparence subsective. Ces adjectifs sont plus dépendants de la sémantique du nom et du contexte que les adjectifs à lecture intersective « simples ». Nous avons posé l'hypothèse que la polydéfinitude, dont la fonction est la sélection d'une entité au sein d'un ensemble d'entités similaires, était considérée par les locuteurs comme pragmatiquement plus appropriée pour les syntagmes présentant de tels adjectifs. La polydéfinitude « pallie » en quelque sorte le fait que ces adjectifs soient plus dépendants du nom que les adjectifs intersectifs simples.

Les adjectifs empruntés en romani d'Albanie ne sont pas pleinement intégrés à la syntaxe héritée, contrairement à l'assertion globale sur l'emprunt des adjectifs par Matras (2009 : 188). On observe les phénomènes suivants :

Tableau 89 : Modèles syntaxiques des adjectifs empruntés par les participants albanais (hors témoins)

	A antéposé en romani (DEF A N)	A postposé en romani (DEF N A + DEF N DEF A)	Total
A antéposé en L2 N occurrences (%)	**fusion** 21 (25 %)	**compartimentation** 63 (75 %)	84 (100 %)
A postposé en L2 N occurrences (%)	**intégration** 0 (0 %)	**convergence + compart.** 18 (100 %)	18 (100 %)
A international	15 (8,29 %)	166 (91,71 %)	181 (100 %)
Total	36 (12,72 %)	247 (87,28 %)	283 (100 %)

Le Tableau 89 montre que les adjectifs empruntés dans les tests sont globalement beaucoup plus postposés (87,28 %) qu'antéposés (12,72 %). Beaucoup d'adjectifs internationaux sont employés, pour lesquels il est difficile de déterminer une origine unique donc une syntaxe précise. Ils sont antéposés dans une minorité de cas (quinze, soit 8,29 %) et postposés dans la très grande majorité des

cas (cent quatre, soit 91,71 %). Ils suivent le principe de *compartimentation* : un adjectif étranger n'est pas placé dans la position canonique, mais dans une place particulière, postposée au nom.

Les adjectifs dont il est possible de déterminer l'origine peuvent être postposés dans la langue d'origine, toujours ou de manière non-marquée : nous n'avons trouvé aucun cas d'*intégration* de tels adjectifs dans la syntaxe héritée (position antéposée). En revanche nous avons trouvé onze cas hybrides, où ces adjectifs postposés dans leur langue d'origine le sont également en romani. En effet ils relèvent à la fois de la *convergence* (puisque postposer l'adjectif permet de satisfaire la syntaxe de la langue de contact) et de la *compartimentation* (puisque l'adjectif n'est pas placé dans la position héritée, antéposée, mais dans la place particulière, postposée au nom).

Les adjectifs dont il est possible de déterminer l'origine peuvent être antéposés dans la langue d'origine, toujours ou de manière non-marquée: nous avons trouvé vingt-et-un cas (35 %) où cette syntaxe de la langue de contact correspond à la syntaxe héritée (position antéposée). Nous qualifions ces cas de *fusion* au sens de Matras (2002 : 211) : il n'est pas possible d'établir une séparation des langues pour cette catégorie syntaxique. Elle semble d'ailleurs restreinte à un certain profil sociolinguistique (variété arli). Au contraire, nous avons trouvé trente-neuf cas (65 %) de *compartimentation*, où l'adjectif à vocation antéposée n'est pas placé dans la position héritée, antéposée, mais dans la place particulière, postposée au nom.

Si la grammaire des lexèmes empruntés est globalement compartimentée par rapport à la grammaire des lexèmes hérités, DEF N DEF A semble une sorte de zone franche, qui présente des taux élevés quels que soient l'origine et le type de l'adjectif. C'est parce qu'elle est la structure la plus appropriée dans le contexte de la tâche, pour exprimer le contraste entre des référents se distinguant par une propriété adjectivale.

Le Tableau 89 a beau présenter une diversité de modèles, on constate le 0 % d'intégration d'adjectifs postposés en antéposition et les 75 % de compartimentation d'adjectifs antéposés vers la postposition (ainsi que les 91,71 % de traitement postposé des adjectifs internationaux). Cela prouve que la syntaxe des adjectifs est bel et bien compartimentée. Matras (2002 : 193, 210-213) décrivait le phénomène sur le plan morphologique : il se situe aussi sur le plan syntaxique. On peut la considérer comme structurelle du romani d'Albanie d'aujourd'hui, parce qu'elle concerne toutes les variétés – du moins dans le contexte énonciatif qui est celui de notre test. La variété arli s'est démarquée comme la seule à pratiquer la convergence – nous étudierons la variable sociale VARIETE en section 3.6 (p. 378) et nous demanderons si les variétés sont stables ou si des changements sont en cours.

La compartimentation concerne également la morphologie. Nous avons trouvé une majorité d'adjectifs invariables, avec aussi une diversité de phénomènes (Tableau 90) :

Tableau 90 : Modèles morphologiques des adjectifs empruntés par les participants albanais

	A variable en romani	A invariable en romani
A variable L contact	**fusion** *zeleno ~ zeleni*	**compartimentation** *žolto ~ *žolti*
A invariable L contact	**intégration** *ješil ~ ješilka*	**convergence + compart.** *saji ~ saji*

Les cas d'invariabilité en romani (*žolto*), alors que l'adjectif est variable dans la langue d'origine montrent que le romani, *compartimente* sa morphologie entre adjectifs empruntés, invariables, et adjectifs hérités, variables. Les cas d'invariabilité en romani (*sarija*), quand l'adjectif serait quoi qu'il arrive invariable dans la langue d'origine (du moins quant à la variation de genre), peuvent être qualifiés d'hybrides. En effet ils relèvent à la fois de la *convergence* (puisque le romani suit le modèle d'invariabilité de l'autre langue) et de la *compartimentation* (puisque le romani compartimente sa morphologie adjectivale, entre adjectifs empruntés, invariables, et adjectifs hérités, variables).

Les cas de variabilité en romani (distinction de genre) montrent que la langue peut intégrer morphologiquement des adjectifs empruntés : ainsi *zeleno*, vert, est-il réanalysé comme un masculin, d'où *zeleni* au féminin. Dans la mesure où le romani satisfait aussi en cela le modèle morphologique de la langue d'origine, qui distingue le masculin du féminin (*zelen/zelena*), on peut parler non seulement d'intégration, mais peut-être aussi de *fusion* : il n'y a pas de séparation des langues pour cette catégorie syntaxique, il y a seulement une différence dans les désinences employées. *Zeleno* est à la fois slave et romani, *zeleni* est un mot slave fléchi à la manière romani.

On peut parler d'intégration dans le cas de *ješil ~ ješilka* également, étant donné que ni l'albanais ni le turc (langues d'emprunt de cet adjectif) ne feraient de distinction de genre dans la configuration qui nous intéresse. *Ješil* est à la fois albanais/turc et romani, *ješilka* est une innovation romani. Il ne s'agit ni d'une convergence ni d'une fusion, puisque la mise au féminin suit un modèle grammatical romani.

Le Tableau 90 a beau présenter une diversité de modèles, les proportions ne sont pas équivalentes : *zeleni* et de *ješilka* représentent des cas uniques et limités à quelques locutrices. La quasi-totalité des occurrences relèvent de la compartimentation et de cas hybrides dus à des langues où l'adjectif est invariable. La compartimentation sur le plan morphologique décrite par Matras (2002 : 193, 210-213) et Adamou et Granqvist (2015) est donc confirmée pour le romani d'Albanie – du moins dans le contexte énonciatif qui est celui de notre test.

Enfin, l'import de l'article de connexion albanais (*e mesme*), constitue lui aussi un cas hybride de convergence et de compartimentation. Nous avons observé des cas où il était perdu (*o lapsi verð*), ce qui est l'indice d'une intégration dans la grammaire romani. Cela peut donner lieu à une variation dans la polydéfinitude, où le deuxième article défini est soit albanais (*o lapsi i verð*), soit ambigu, soit romani (*i bojàva i verða*).

Le Tableau 71 p. 329, répété ici et dressé à partir de Matras (2009 : 187-190) présentait une typologie de l'emprunt adjectival dans les langues du monde, en fonction de la syntaxe et de la morphologie. Nous pouvons ajouter, en rouge, la ligne du romani d'Albanie :

Tableau 71 : Typologie de l'emprunt adjectival suivant (Matras 2009 : 187-190)

	Syntaxe	Morphologie
allemand	✓: antéposition	✓: flexion héritée
hébreu	✓: postposition	X : assignation à une classe flexionnelle particulière
yiddish	✓: antéposition	✓: flexion héritée (adaptée si besoin)
romani **romani albanais**	? ✓: antéposition (Burgenland) X : postposition	✓: flexion héritée (adaptée si besoin) X : absence de morphologie (Burgenland) X : absence de morphologie
persan, kurde	✓: postposition	✓: absence de flexion
turc	✓: antéposition	✓: absence de flexion
ourdou	?	X : absence de morphologie
anglais	X : postposition	X : flexion de la langue d'origine
maltais	✓: postposition (héritée & empruntée)	✓: flexion héritée ou partiellement empruntée X : absence de flexion

Le romani d'Albanie est l'inverse du type anglais : tandis que ce dernier n'intègre pas du tout l'adjectif pour conserver au mieux la syntaxe et la morphologie de la langue d'origine, le romani n'intègre pas du tout l'adjectif, pour lui assigner une syntaxe et une morphologie particulière.

L'état des lieux chez les témoins présente des similarités : on trouve une très grande majorité d'adjectifs hérités et une minorité d'adjectifs empruntés. Les adjectifs hérités sont majoritaires en DEF A N… mais minoritaires en DEF N A. L'échelle d'héritage semble donc respectée à la lettre chez les témoins, ce qui peut être en réalité causé par la quasi absence de la polydéfinitude, réduite à un simple DEF A N > DEF N A. La polydéfinitude ne peut jouer le rôle de « zone franche » qu'elle joue dans les variétés albanaises. Puisque DEF A N représente l'écrasante majorité des syntagmes de la tâche, on peut dire que la fonction de contraste est assumée par elle seulement : le lieu syntaxique pour exprimer le contraste dans ces trois variétés est le *prenominal slot*. Une autre preuve en est que les adjectifs faussement subsectifs (taille et autres), sont tous employés en structure DEF A N (à l'exception d'un adjectif de taille et d'un adjectif autre, d'ailleurs emprunté).

Les rares cas de postposition de l'adjectif, en syntagme mono- ou polydéfini, s'expliquent presque tous par le contact de langues, plus précisément par le principe de compartimentation syntaxique. Parmi les six syntagmes DEF N DEF A et DEF N A, on trouve en effet quatre adjectifs empruntés et seulement deux exceptions avec adjectif hérité (*baro, parno*).

Les adjectifs empruntés sont, comme dans les variétés albanaises, presque tous des adjectifs de couleur (intersectifs). Les adjectifs empruntés suivent le modèle de fusion syntaxique dans 12 cas sur 16, c'est-à-dire qu'ils proviennent de langues à adjectif antéposé, et restent antéposés en romani. Les quatre autres sont postposés, suivant par là le modèle de compartimentation syntaxique. Morphologiquement, nous n'avons pu détecter qu'un cas de compartimentation et un cas de fusion.

2.5 Synthèse des variables linguistiques

Les variables linguistiques nous ont permis d'étudier en profondeur comment se construit la polydéfinitude en romani d'Albanie :

- Elle se construit exclusivement avec l'article défini, si l'on excepte un cas de répétition du démonstratif, mais jamais avec l'article indéfini. (DEFINITUDE)
- Elle se construit quasiment exclusivement avec un adjectif postposé, si l'on excepte trois cas d'antéposition. (ORDRE DES CONSTITUANTS)
- Nous avons constaté que la postposition de l'adjectif est non-marquée dans le contexte des tests. Il s'agit d'une rupture avec les autres contextes énonciatifs, voire avec les autres variétés, pour lesquels la position non-marquée de l'adjectif est antéposée.
- Lorsque plusieurs adjectifs encadrent le nom, c'est l'adjectif postposé qui porte le contraste, ce qui confirme que la position non-marquée de l'adjectif à visée contrastive (contexte de la tâche) est postposée.
- La polydéfinitude donne lieu à variation morphologique. Nous avons observé une répétition des marques flexionnelles et une répétition de la préposition, lorsqu'elle a fusionné avec l'article défini.

Le placement des adjectifs sur l'échelle d'héritage (ou « héritéité ») en fonction de leur origine laisse à penser que la structure la plus innovative est DEF N A et que DEF N DEF A constitue elle aussi une innovation, pour ainsi dire plus ancienne. Une telle hypothèse demande à être étayée : nous y reviendrons en section 2.5 du chapitre VI, p. 457.

La rupture avec les autres variétés ne s'établit pas par la comparaison avec la littérature sur les autres variétés ou sur le romani en général, mais aussi par la comparaison avec les **témoins**. On trouve chez eux une syntaxe très différente :

- la polydéfinitude est quasiment inexistante (la construction est cependant la même : deux articles définis, adjectif postposé). (DEFINITUDE)
- la postposition de l'adjectif est quasiment inexistante. (ORDRE DES CONSTITUANTS)
- L'antéposition de l'adjectif est non-marquée dans le contexte de la tâche.
- Lorsque plusieurs adjectifs encadrent le nom, c'est l'adjectif antéposé qui porte le contraste, ce qui confirme que la position non-marquée de l'adjectif à visée contrastive est antéposée.

L'analyse de la polydéfinitude nous a permis de déterminer la fonction des différentes structures syntagmatiques dans le contexte de la tâche :

- DEF A N est la structure privilégiée des adjectifs hérités intersectifs (aspect)
- DEF N A est la structure privilégiée des adjectifs empruntés intersectifs (couleur)
- DEF N DEF A est privilégiée à part égale par les adjectifs de toutes origines. Elle est en revanche la structure privilégiée des adjectifs à lecture intersective en apparence subsective, qui se trouve être hérités (taille, *autre*).

Les témoins macédoniens et kosovar serbe suivent le schéma albanais sur ces points :

- DEF A N est la structure privilégiée des adjectifs hérités.
- DEF N (DEF) A est la structure privilégiée des adjectifs empruntés.

Pour comprendre la polydéfinitude, on voit qu'il faut donc combiner une explication par l'origine et une explication par la sémantique de l'adjectif (TYPE D'ADJECTIF).

Les adjectifs empruntés du romani d'Albanie ne sont pas intégrés à la syntaxe héritée puisque l'on n'a pu recenser aucun cas d'intégration. On a observé quelques cas hybrides relevant à la fois de la convergence et de la compartimentation, quelques cas de fusion et une grande majorité de cas de compartimentation. Ils ne sont pas non plus intégrés à la morphologie héritée puisque la très grande majorité d'entre eux restent invariables, relevant soit de la compartimentation, soit de l'hybridité convergence + compartimentation. Chez les témoins en revanche, la fusion syntaxique semble la stratégie majoritaire en cas d'emprunt d'un adjectif, ce qui est rendu possible par le contact dominant avec des langues slaves. La compartimentation représente le reste des cas.

Les variétés romani d'Albanie, dans notre test, se caractérisent donc par la compartimentation tant sur les plans morphologique (ce que laissait prévoir la littérature) que syntaxique (ce qui n'était pas prévu). Cette compartimentation rappelle dans une certaine mesure le type anglais de la typologie établie à partir de Matras (2009 : 187-190) – dans une certaine mesure seulement puisque le romani, en compartimentant, évite en fait d'importer la syntaxe et la morphologie de la langue d'origine, mais crée une grammaire empruntée spécifique, indifférente à la langue d'emprunt. Cette stratégie est cognitivement plus économique qu'une stratégie telle que celle de l'anglais, qui requiert de connaître la syntaxe et la morphologie des adjectifs dans les différentes langues d'emprunt.

Nous avons observé que l'adjectif *vaver*, autre, fait l'objet d'un usage particulier :

- Il est parfois utilisé, en position antéposée, en sus d'un adjectif postposé de couleur ou d'aspect… qui serait déjà suffisant pour identifier le référent (exemple (212) de structure DEF A N A, et, dans une moindre mesure, exemple (213); exemples (227) et (228) de structure DEF A N DEF A).
- En position postposée, où il suffit à l'identification, les locuteurs peuvent toutefois éprouver le besoin d'ajouter d'autres adjectifs après lui (exemple (229) de structure DEF N A/ DEF A ; exemple (230) de structure DEF N/ DEF A/ A ; exemples (232), (234) et (235) de structure DEF N

DEF A DEF A). Les trois derniers exemples cités sélectionnent notamment un référent qui contraste de manière binaire avec un autre.

Vaver, qui n'a pas d'autre sens que le contraste, est le plus approprié, voire cognitivement le plus économique, pour identifier le référent. Pourquoi, alors, ajouter du matériel linguistique ? De ces deux constats, on peut tirer la conclusion que l'adjectif *vaver*, autre, a une sémantique plus faible que les autres adjectifs. C'est qu'il ne transmet pas une propriété du référent, mais une propriété de l'actualisation du nom en discours. Dès lors que le « stock » de référents contient plus de deux référents possibles, ce qui est le cas avec les objets de la tâche de localisation, il ne suffit pas à opérer une sélection efficace. Le corpus spontané permet d'observer l'usage de l'adjectif lorsque l'identification précise n'est pas absolument nécessaire. C'est le cas de l'exemple (210), où la locutrice ne permet pas l'identification précise du référent : il s'agit d'une mère de quatre filles qui mentionne son « autre fille ». Le stock contient quatre référents possibles, dont deux sont exclus par la situation d'énonciation (une fille est présente lors de l'enregistrement) et par la connaissance générale (une deuxième fille a déjà été mentionnée en discours), mais reste l'ambiguïté entre les deux derniers possibles. Lorsque le stock de référents contient seulement deux référents, la sélection est possible avec *vaver*. C'est le cas de l'exemple (239) : le locuteur a deux frères, il mentionne d'abord « le grand frère » puis « l'autre frère ». C'est également le cas de l'exemple (241) : le locuteur parle des « plus anciens Roms » d'Albanie, puis des « autres Roms ».

3. Résultats : variables sociales

3.1 Introduction

3.1.1 Variantes étudiées

Nous recensons dans cette section le nombre des syntagmes nominaux suivants :

- polydéfinis (DEF N DEF A)
- monodéfinis avec A postposé (DEF N A)
- monodéfinis avec A antéposé (DEF A N)
- poly- ou monodéfinis complexes
- présentant un A postposé quel que soit le déterminant
- présentant un A antéposé quel que soit le déterminant

Les pourcentages de présence sont calculés par rapport à chaque locuteur pris individuellement. En effet, nous ne pouvons travailler à partir du nombre absolu d'occurrences, puisque les données sont inégales selon les personnes : temps de parole/temps d'enregistrement, nombre de passages des tests. C'est donc la proportion d'apparition et de non-apparition des structures les unes par rapport aux autres qui nous intéresse.

3.1.2 Variables étudiées

La structure du syntagme nominal et l'ordre des constituants sont étudiés en fonction des variables suivantes, qui présentent chacune différentes *modalités*, c'est-à-dire peuvent prendre différentes valeurs (Tableau 91). On a vu en section 1.1.1 du chapitre I, p. 5, qu'il est difficile de fournir une idée ne serait-ce qu'approximative de la réalité démographique des Roms en Albanie, quant au nombre total de personnes, à la pyramide des âges selon l'âge et selon le genre, la répartition par groupes, etc.

Tableau 91 : Récapitulatif des variables sociales, des modalités, et de l'échantillon étudié

Variable	Modalités possibles	Test 1	Test 2	Témoins	Commentaires
GENRE	0 = masculin	6	15	2	
	1 = féminin	8	16	1	
CLASSE D'AGE	0 – moins 15 ans	3	2	0	Locuteurs non mariés, qui ne travaillent pas et sont considérés comme des enfants par les Roms et par la société dominante.
	1 – 15 – 39 ans	7	16	2	Locuteurs dans la force de l'âge, dont un grand nombre sont mariés, ont des enfants et travaillent. Ils sont considérés comme de jeunes adultes par les Roms.
	2 – 40 – 59 ans	3	8	1	Locuteurs dans la force de l'âge, mariés et avec des enfants, qui travaillent, mais dont un grand nombre sont déjà grands-parents. Ils sont considérés comme des adultes par les Roms.
	3 – plus 60 ans	1	5	0	Locuteurs mariés, qui ne travaillent pas et sont considérés comme des seniors par les Roms et par la société dominante.
EDUCATION	0 – 0 à 5 années de scolarité	3	11	0	Locuteurs qui n'ont pas reçu de formation du tout ou ont reçu une formation primaire du cycle inférieur (la première moitié du cycle d'école obligatoire). Nous incluons ici les personnes ayant réalisé 5 années d'école, considérant que cette 5e année peut correspondre soit à une année de décrochage scolaire, soit à une année de redoublement, soit à une année de deuxième partie du cycle primaire, mais que le niveau acquis reste celui du cycle primaire accompli (cf. *infra* et et Tableau 2, p. 9).
	1 – 6 à 12 années de scolarité	7	13	1	Locuteurs qui ont reçu la formation primaire du cycle inférieur et du cycle supérieur, ainsi que ceux qui ont fait le lycée, quel que soit le lycée (cf. *infra* et Tableau 2, p. 9).
	2 – plus de 12 années de scolarité	4	7	2	Locuteurs qui ont suivi une formation supérieure (cf. *infra* et Tableau 2, p. 9).
MILITANTIME	0 – non-militant	11	24	3	cf. *infra* et Tableau 92
	1 – militant	3	7	0	cf. *infra* et Tableau 92
VARIETE	0 – néo-locuteur arli	1	1	0	cf. section 2.4.6. du chapitre I, p. 42
	1 – mečkar	4	6	0	
	2 – arli	6	12	0	
	3 – čergar I	0	2	0	

Variable	Modalités possibles	Test 1	Test 2	Témoins	Commentaires
	4 – arli & čergar I / čergar I & arli	3	7	0	
	5 – arli & bamlija	0	1	0	
	6 – mečkar & čergar I	0	1	0	
	7 – mečkar & bamlija	0	1	0	
	8 – kovački	0	0	1	
	9 – džambazo	0	0	1	
	10 – xoraxane	0	0	1	

3.1.3 Le genre

Le genre n'est pas assimilable au seul sexe biologique. Le genre est une identité sociale : pour cette raison, elle émerge dans la façon dont nous interagissons avec autrui (Meyerhoff 2011 : 232) et dont nous négocions notre place dans le groupe. Nous considérons le genre comme une construction sociale et non comme un simple reflet du sexe biologique : « *another advantage to seeing gender as a socially complex category, and not simply as a reflex of biological sex, is that it makes the question of how gender interacts with other social identities a central concern, rather than a tangential (or side) matter* » (Meyerhoff 2011 : 232).

Meyerhoff (2011 : 280) montre qu'un principe stable en sociolinguistique veut que les femmes utilisent le standard plus que les hommes. En effet, les recherches montrent que les femmes auraient une plus grande sensibilité à la différence entre traits (ou langue) standard et non-standard, et une meilleure aptitude à passer de l'un à l'autre (Meyerhoff 2011 : 281). Elles ont également une meilleure sensibilité à ce qui est prohibé par la norme et donc une meilleure aptitude à l'éviter. Cela peut être dû au fait que les hommes sont jugés plutôt quant à leurs actions, tandis que les femmes sont jugées plutôt quant à leur apparence. Cela amène les femmes à accorder plus d'attention à leur façon de s'exprimer et notamment aux marqueurs stylistiques de leur discours. Elles feraient ainsi plus recours aux ressources symboliques en général (discours, vêtements, maquillage) afin d'établir et d'affirmer leur position (ou leur opposition) et leur identification dans un groupe social (Meyerhoff 2011 : 281, Eckert 2000). Chambers (2013 : 302) explique que « *within the class norms, sociolinguistic research has discovered that women use fewer stigmatized and nonstandard variants than do men of the same social group in the same circumstances.* […] *Consequently, speech communities are marked by consistent and, as we shall see, partly predictable linguistic correlations with sex* ». Cela ne signifie pas que les femmes soient nécessairement conservatrices sur le plan linguistique. Elles auraient simplement une meilleure aptitude à détecter toute nouvelle variante émanant d'une forme prestigieuse de la langue. Cela les rend porteuses d'innovation lorsqu'elles ont conscience d'un changement en cours, et que ce changement s'effectue au moyen d'une variante jugée positivement par la communauté (Meyerhoff 2011 : 282).

Il existe toutefois des exceptions à ces principes, dépendant notamment de facteurs sociaux. En effet, il faut se garder d'essentialiser le genre, et d'associer telle ou telle variante, tel ou tel comportement à un genre particulier. C'est le rôle social que joue un locuteur, en tant qu'homme ou femme, ainsi que le réseau relationnel dans lequel il est impliqué, qui provoque la distribution genrée des variantes linguistiques (Meyerhoff 2011 : 293). Le genre n'est donc pas une variable monolithique.

3.1.4 Le système scolaire en Albanie

Le système scolaire comporte les étapes décrites dans le Tableau 2, p. 9, répété ici :

Tableau 2 : Organigramme du système scolaire albanais

<table>
<tr><th>Cycle</th><th>Nom</th><th>Durée</th><th>Age</th><th>Statut</th></tr>
<tr><td>pré-primaire</td><td>parashkolla (« pré-école »)</td><td>3 ans</td><td>3 – 6 ans</td><td>facultatif</td></tr>
<tr><td rowspan="2">primaire</td><td>shkollë tetëvjeçare – cikël i ulët (« école de 8 ans – cycle inférieur »)</td><td>4 ans</td><td>6 – 10 ans</td><td rowspan="2">obligatoire</td></tr>
<tr><td>shkollë tetëvjeçare – cikël i lartë (« école de 8 ans – cycle supérieur »)</td><td>4 ans</td><td>10 – 14 ans</td></tr>
<tr><td colspan="5">Passage du provime lirimi (« examen de fin d'études »), en vue du dëftesë lirimi (« certificat de libération »)</td></tr>
<tr><td rowspan="3">secondaire</td><td>shkollë e mesme e përgjithshme (« école moyenne générale »), communément appelée gjimnaz (« lycée »)</td><td rowspan="3">3, 4 ou 5 ans</td><td>14 – 18 ans</td><td rowspan="3">facultatif</td></tr>
<tr><td>shkollë e mesme profesionale (« école moyenne professionnelle »)</td><td>14 – 17 ans</td></tr>
<tr><td>shkollë teknike (« école technique »)</td><td>14 – 19 ans</td></tr>
<tr><td colspan="5">Passage du baccalauréat, dëftesë pjekurie (« certificat de maturité »), communément appelé matura</td></tr>
<tr><td>tertiaire</td><td>universitet (« université »)
institute i lartë (« institut supérieur »)
akademi (« académie »)</td><td>2 à 6 ans</td><td>± 18 ans – ± 24 ans</td><td>facultatif</td></tr>
<tr><td colspan="5">Passage d'un diplôme, diplomë : dëshmi specializimi (« certificat de spécialisation ») ou doktor (« docteur »)</td></tr>
</table>

3.1.5 Le militantisme

Acuña et Adamou (2013) ont constaté que l'engagement en tant que militant de son peuple et de sa culture peut avoir une influence sur la grammaire idiosyncrasique : l'innovation qui fait l'objet de leur étude (l'usage du clitique sujet *lo, la, le* pour répliquer en romani la copule espagnole *estar*) est présente chez tous les locuteurs... sauf chez les militants de la cause rom. En effet, il s'avère que ces derniers ont une idéologie linguistique, mais aussi des réseaux sociaux qui peuvent être différents. Acuña et Adamou (2013) parlent même de « rapport professionnel » à la langue. Comme un certain nombre des participants à la tâche ont ou pourraient avoir un tel rapport au romani, nous avons décidé d'établir une variable MILITANTISME. Comment la définir ? Il existe différents types d'engagement personnel chez les locuteurs :

- engagement socio-politique en faveur des droits des Roms, contre l'exclusion, le racisme, incluant la participation à des manifestations et formations diverses sur ces thématiques.
- engagement spécifique en faveur de la culture rom (traditions, histoire, identité) et particulièrement de la langue romani, incluant la mise en place ou la participation à des cours de langue (souvent bénévoles).

Ces deux types de militantisme, que nous qualifierons de social et de linguistique, peuvent être réalisés séparément, ensemble ou bien se succéder au cours de la vie des locuteurs. Le type d'engagement qui est le plus susceptible d'avoir un impact sur la manière de s'exprimer est l'engagement linguistique car il peut modifier l'idéologie linguistique par une attitude volontariste (par exemple « puriste »).

Ils coexistent au niveau local et/ou régional, au niveau national et au niveau européen et/ou international. À notre connaissance, les Roms des différents groupes d'Albanie sont en contact fréquent entre eux (par des liens de voisinage, d'amitié, de famille, de travail...), et donc habitués à « entendre » leurs variétés respectives. Ainsi peut-on avancer que l'effet du militantisme aux niveaux local, régional et national a globalement le même impact, dans la mesure où il consiste à fréquenter locuteurs de variétés différentes mais déjà connues, ayant toutes pour langue de contact dominante l'albanais. Au contraire, le militantisme à l'échelle internationale fait entendre au locuteur d'autres variétés, issues d'autres pays, avec d'autres traditions, en contact avec d'autres langues : c'est le phénomène d'émergence de la *khetani ćhib*, langue de rassemblement (litt. langue commune) (cf. section 2.2.2, p. 35).

Pour cette étude particulière, nous considérons que deux types d'engagement peuvent avoir un impact sur la grammaire : engagement linguistique, engagement au niveau européen et/ou international. Nous choisissons de ne pas faire entrer dans cette catégorie une personne ayant par exemple milité au niveau national en Albanie puis au niveau national en Italie, car cet engagement lui a permis de construire un réseau militant national albanais, puis national italien, mais pas nécessairement romani transnational ni plurinational. Nous décidons de ne pas tenir compte de l'époque de l'engagement en question (passé ou présent), estimant que l'influence est soit établie soit en train de s'établir.

Le Tableau 92 présente le type d'engagement des locuteurs ayant participé aux tests (Albanais et témoins).

Tableau 92 : L'engagement militant de l'ensemble des participants aux tests 1 et 2

	Social	Linguistique	Échelle nationale	Échelle internationale
Engagé	20	4	20	6
Non-engagé	17	33	17	31
Total	37	37	37	37

3.1.6 Les variétés

Etant accueillie par une famille arli, dans une ville où ils représentent l'immense majorité des Roms, nous avons eu accès à une majorité de locuteurs de l'arli, ce qui a rendu difficile l'obtention d'un équilibre entre les différentes variétés.

N'ont été classées comme locutrices de telle variété que les personnes dont la famille parle en majorité cette variété. Les autres sont au contraire issues de couples « mixtes », dont les parents (et donc les deux parties de la famille élargie) sont/étaient locuteurs de deux variétés différentes. Nous postulons que ces personnes ont une connaissance au minimum passive de deux variétés dans le cadre familial. Il peut aussi arriver que la personne ait été élevée ou sociabilisée dans une variété autre que celle de ses parents. Dans ces cas où il n'y a pas de variété de référence unique, il semble difficile et inutile de juger si l'une d'elles est dominante et laquelle, dans quel domaine linguistique, à quel degré

et selon quels critères. En outre, s'il nous a été possible, pour certaines personnes, de savoir si l'une des deux variétés est dominante et laquelle, ce n'est pas le cas de toutes. Nous avons donc renoncé à juger et préféré créer des catégories spécifiques pour ces locuteurs « à deux variétés » (arli & čergar I / čergar I & arli, arli & bamlija, mečkar & čergar I, mečkar & bamlija).

Un locuteur de famille arli a en outre déclaré ne pas avoir appris la langue en famille, mais en tant qu'adulte, à l'occasion de cours de langue spécifiques. Nous avons choisi de lui consacrer une catégorie particulière, « néo-locuteur arli » (*new speaker*, voir la section 2.4.6, p. 42).

Enfin, trois personnes non albanaises ont accepté de participer au test et ont donc pu servir de groupe-témoin. Elles ont nommé leurs variétés *xoraxane* (du Kosovo), *kovački* et *džambazo* (de Macédoine du Nord).

Le **xoraxane** et le **džambazo** appartiennent au groupe *vlax* du sud. Friedman (2017 : 34) mentionne le *džambaz* macédonien dans son étude sur les dialectes arli de Skopje. Le terme provient directement du turc *cambaz*, maquignon, acrobate, saltimbanque. Ce groupe a été le dernier à s'installer en ville, en raison de sa profession itinérante.

Le **kovački** appartient au groupe balkanique II / *zis* (cf. section 2.1.3, p. 32). Friedman (2017) mentionne le kovač (ou burgudži) macédonien dans la même étude. Le terme *burgurdži* provient du turc *burgucu*, artisan qui fabrique des chignoles. Ce groupe est connu également sous le nom *esnaf*, guilde, membre d'une corporation, propriétaire d'un atelier. Ils seraient arrivés en Macédoine depuis le nord-est de la Bulgarie (Matras 2002 : 223, Boretzky et Igla 2004, Friedman 2017 : 34). La variété de *bugurdži* (métathèse de *burgurdži*) kosovare décrite par Boretzky (1993) n'en est probablement pas éloignée. Le terme *kovač* est un nom slave signifiant forgeron. À Skopje, ce groupe porte d'autres noms : *kalajdži*, étameur (terme du turc roumélique occidental), *arabadži*, charretier (même variété de turc) ; en Bulgarie : *drəndar*, cardeur de laine ; au Kosovo : *rabadži*, charrettier et *kovač*, forgeron.

3.2 Variable 1 : GENRE

3.2.1 Participants albanais

3.2.1.1 Structure des syntagmes nominaux définis

La distribution des syntagmes nominaux définis incluant un adjectif épithète en fonction du GENRE est représentée en Figure 49 (cf. Tableau 130).

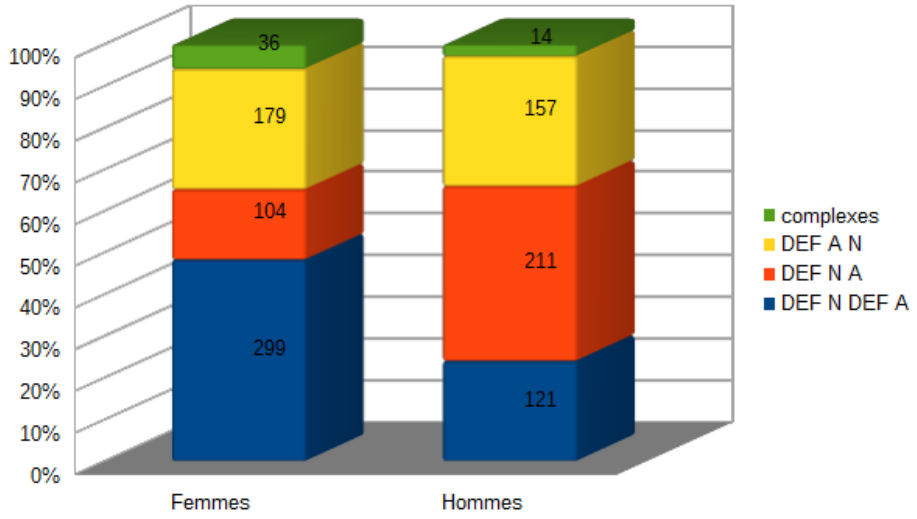


Figure 49 : Structure des SN définis produits par les participants des tests 1 et 2 en fonction du GENRE

On observe que les femmes produisent plus de syntagmes polydéfinis DEF N DEF A que les hommes – 299 occurrences pour 121, soit 48,38 % et 24,06 % de leur production respective. 71,19 % des polydéfinis DEF N DEF A est donc produit par les femmes, et seulement 28,81 % par des hommes. Parmi les monodéfinis, que les femmes produisent donc en proportion plus restreinte, on observe une part nettement marquée pour l'antéposition de l'adjectif (179 occurrences pour 104, soit 28,96 % de DEF A N pour 16,83 % de DEF N A sur le total des syntagmes), tandis que les hommes, qui produisent donc proportionnellement plus de monodéfinis que les femmes, présentent une tendance nettement marquée à la postposition de l'adjectif (211 occurrences soit 41,95 %) plutôt qu'à l'antéposition (157 occurrences soit 31,21 %) de celui-ci. Les femmes produisent plus de structures complexes (36 soit 5,83 %) que les hommes (14 soit 2,78 %). D'une manière générale, les femmes utilisent autant de structures mono- que polydéfinies, ce qui suppose qu'elles postposent en réalité plus l'adjectif qu'elles ne l'antéposent – les hommes arrivent à une proportion semblable de postposition, mais en usant de la stratégie monodéfinie plutôt que polydéfinie.

3.2.1.2 Proportion d'adjectifs anté- et postposés

La distribution des syntagmes nominaux définis présentant un adjectif épithète anté- ou postposé en fonction du GENRE est représentée en Figure 50 (cf. Tableau 131).

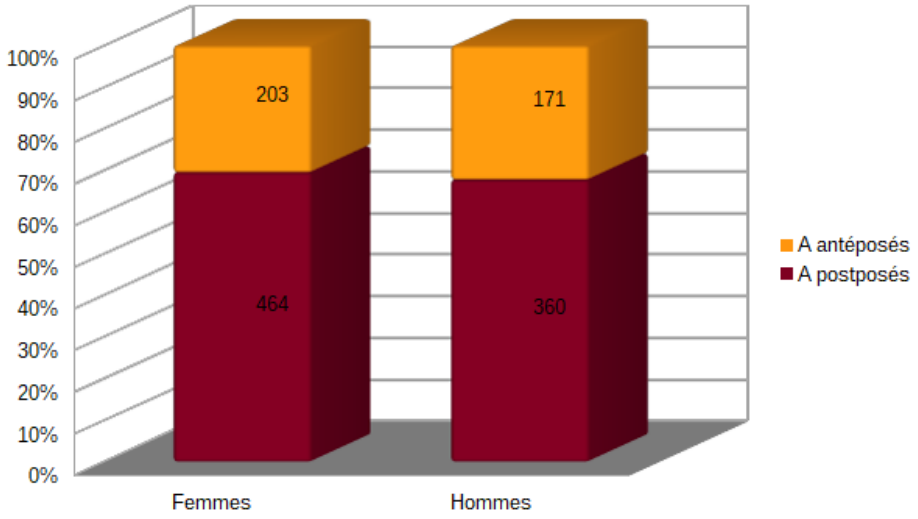


Figure 50 : Proportion d'adjectifs anté- et postposés produits par les participants des tests 1 et 2 en fonction du GENRE

En incluant les autres déterminants dans le calcul on aboutit au même résultat que précédemment quant à la position de l'adjectif. Dans la mesure où la part finale des adjectifs anté- et postposés est sensiblement la même chez les femmes (464 occurrences soit 69,57 %) et chez les hommes (360 occurrences soit 67,80 %), on peut émettre l'hypothèse que la polydéfinitude avec postposition de l'adjectif joue chez les femmes un rôle semblable à celui que joue la simple postposition de l'adjectif chez les hommes.

3.2.2 Témoins

3.2.2.1 Structure des syntagmes nominaux définis

La distribution des syntagmes nominaux définis incluant un adjectif épithète en fonction du GENRE est représentée en Figure 51 (cf. Tableau 132).

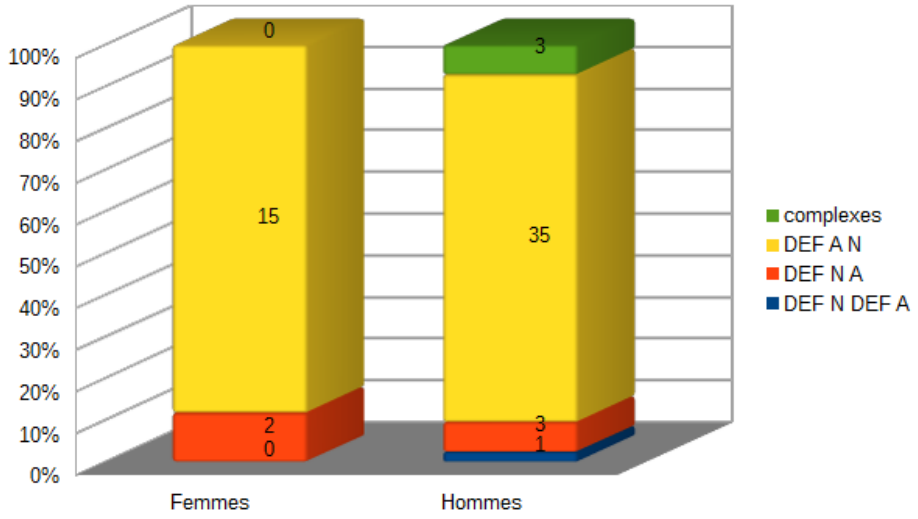


Figure 51 : Structure des SN définis produits par les témoins en fonction du GENRE

Les témoins ont un comportement syntaxique peu différent en fonction du genre : un homme produit un seul syntagme DEF N DEF A, tandis que l'autre homme et la femme n'en produisent aucun. Les trois produisent entre 84 % et 88 % de DEF A N. La femme seule produit autant de DEF N A que les deux hommes ensemble.

3.2.2.2 Proportion d'adjectifs anté- et postposés

La distribution des syntagmes nominaux définis présentant un adjectif épithète anté- ou postposé en fonction du GENRE est représentée en Figure 52 (cf. Tableau 133).

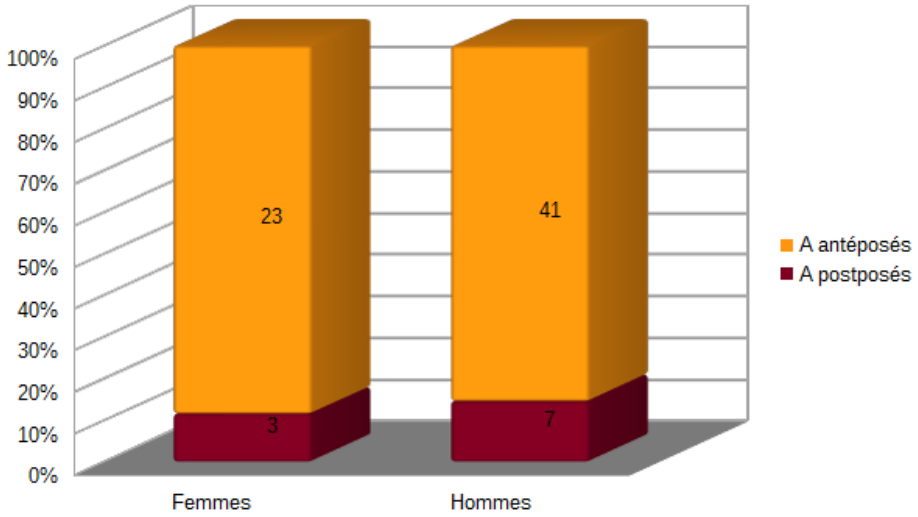


Figure 52 : Proportion d'adjectifs anté- et postposés produits par les témoins en fonction du GENRE

Ils produisent chacun 3 adjectifs postposés avec d'autres déterminants, soit environ 11 à 13 % du total – tous les autres syntagmes présentent un adjectif antéposé. La femme seule postpose globalement autant ses adjectifs que les deux hommes ensemble.

Les deux hommes n'ont pas un comportement syntaxique semblable qui se distinguerait de celui de la femme.

3.3 Variable 2 : CLASSE D'AGE

3.3.1 Participants albanais

3.3.1.1 Structure des syntagmes nominaux définis

La distribution des syntagmes nominaux définis incluant un adjectif épithète en fonction de la CLASSE D'AGE est représentée en Figure 53 (cf. Tableau 134).

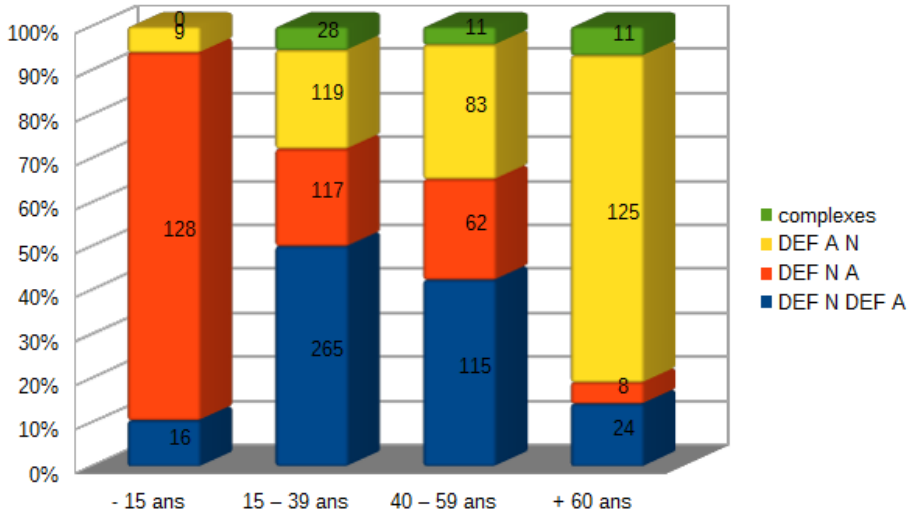


Figure 53 : Structure des SN définis produits par les participants des tests 1 et 2 en fonction de la CLASSE D'AGE

On observe que, dans l'ensemble, les syntagmes DEF N DEF A sont l'œuvre des deux classes d'âge centrales, étant pratiquées en priorité par les 15-39 ans (50,09 %), puis les 40-59 ans (42,43 %), mais moins par les plus de 60 ans (14,29 %) et par les moins de 15 ans (10,46 %). Les classes d'âge extrêmes utilisent de manière presque catégorique la monodéfinitude (89,54 % pour les moins de 15 ans, 79,17 % pour les plus de 60 ans) – les jeunes préférant postposer l'adjectif, les anciens l'antéposer. Les classes d'âge centrales pratiquent un usage plus équilibré des deux structures DEF A N (environ 22 % pour les deux classes d'âges) et DEF N A, avec une préférence pour cette dernière

(22,50 % chez les 15-39 ans et 30,63 % chez les 40-59 ans). D'une manière générale, les classes d'âge centrales utilisent autant de structures mono- que polydéfinies, ce qui suppose qu'elles postposent en réalité plus l'adjectif qu'elles ne l'antéposent.

3.3.1.2 Proportion d'adjectifs anté- et postposés

La distribution des syntagmes nominaux définis présentant un adjectif épithète anté- ou postposé en fonction de la CLASSE D'AGE est représentée en Figure 54 (cf. Tableau 135).

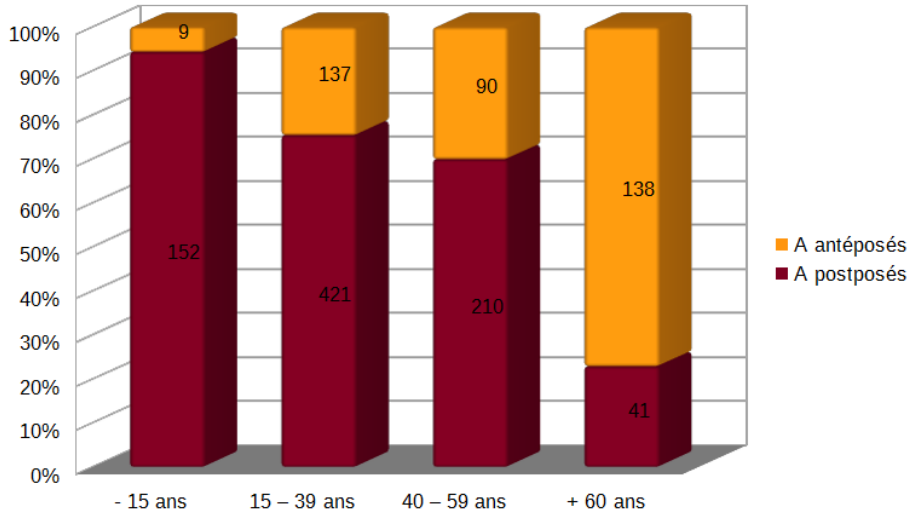


Figure 54 : Proportion d'adjectifs anté- et postposés produits par les participants des tests 1 et 2 en fonction de la CLASSE D'AGE

La postposition de l'adjectif semble être globalement une innovation récente, si l'on en croit sa présence décroissante avec l'âge (mono- et polydéfinis, tous déterminants inclus) : 94,41 %, 75,45 %, 70 % et 22,91 % par opposition à l'antéposition de l'adjectif.

3.3.2 Témoins

3.3.2.1 Structure des syntagmes nominaux définis

La distribution des syntagmes nominaux définis incluant un adjectif épithète en fonction de la CLASSE D'AGE est représentée en Figure 55 (cf. Tableau 136).

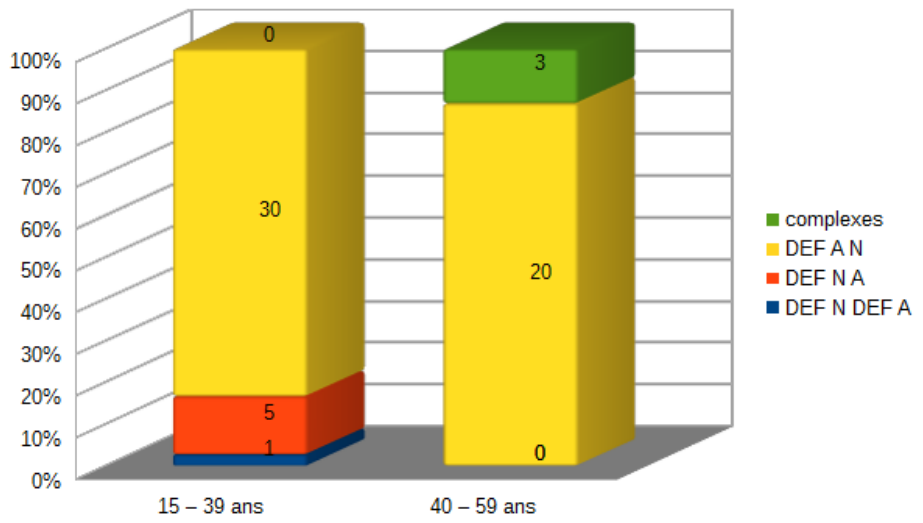


Figure 55 : Structure des SN définis produits par les témoins en fonction de la CLASSE D'AGE

Les témoins n'appartiennent qu'à deux classes d'âge, entre 15 et 59 ans. Ils ont un comportement syntaxique peu différent en fonction de cette variable : un 15-39 ans produit 1 syntagme polydéfini DEF N DEF A, tandis que l'autre locutrice de cette classe et le 40-59 ans n'en produisent aucun. Les deux plus jeunes sont les seuls à postposer l'adjectif (DEF N A ou DEF N DEF A). Les deux personnes les plus jeunes sont les seules à postposer l'adjectif (DEF N A ou DEF N DEF A). Les trois produisent entre 84 % et 88 % de monodéfinis avec adjectifs antéposé DEF A N.

3.3.2.2 Proportion d'adjectifs anté- et postposés

La distribution des syntagmes nominaux définis présentant un adjectif épithète anté- ou postposé en fonction de la CLASSE D'AGE est représentée en Figure 56 (cf. Tableau 137).

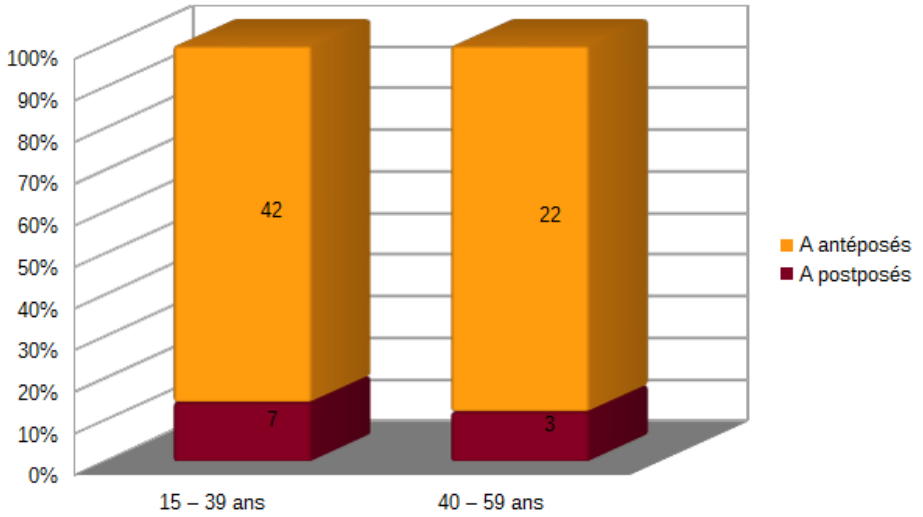


Figure 56 : Proportion d'adjectifs anté- et postposés produits par les témoins en fonction de la CLASSE D'AGE

Ils produisent chacun 3 adjectifs postposés avec d'autres déterminants, soit environ 11 à 13 % seulement – tous les autres syntagmes présentent un adjectif antéposé. La personne la plus vieille postpose à elle seule autant ses adjectifs que les deux plus jeunes qu'elle.

Les deux personnes entre 15 et 39 ans ont donc un comportement semblable légèrement distinct de la personne entre 40 et 59 ans.

3.4 Variable 3 : EDUCATION

3.4.1 Participants albanais

3.4.1.1 Structure des syntagmes nominaux définis

La distribution des syntagmes nominaux définis incluant un adjectif épithète en fonction de l'EDUCATION est représentée en Figure 57 (cf. Tableau 138).

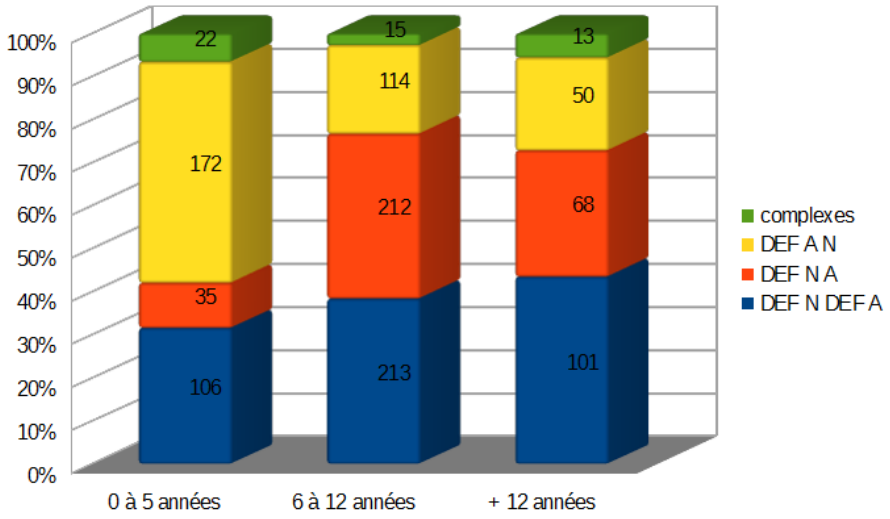


Figure 57 : Structure des SN définis produits par les participants des tests 1 et 2 en fonction de l'EDUCATION

La production de structures DEF N DEF A semble aller croissant avec le niveau d'éducation : 31,64 %, 38,45 % et 43,54 %. Les syntagmes monodéfinis à adjectif postposé DEF N A sont le fait des deux catégories les plus éduquées, en majeure partie par la catégorie centrale. Les deux catégories les plus éduquées utilisent dans la grande majorité des cas ces deux structures DEF N DEF A et DEF N A, ce qui implique qu'elles postposent l'adjectif dans la grande majorité des cas (76,71 % et 72,85 %). Au contraire, les syntagmes monodéfinis à adjectif antéposé DEF A N sont le fait de la catégorie de

locuteurs la moins éduquée, et ce très nettement : 51,34 % d'adjectifs antéposés chez ces locuteurs contre environ 20,58 % et 21,55 % chez les autres.

3.4.1.2 Proportion d'adjectifs anté- et postposés

La distribution des syntagmes nominaux définis présentant un adjectif épithète anté- ou postposé en fonction de l'EDUCATION est représentée en Figure 58 (cf. Tableau 139).

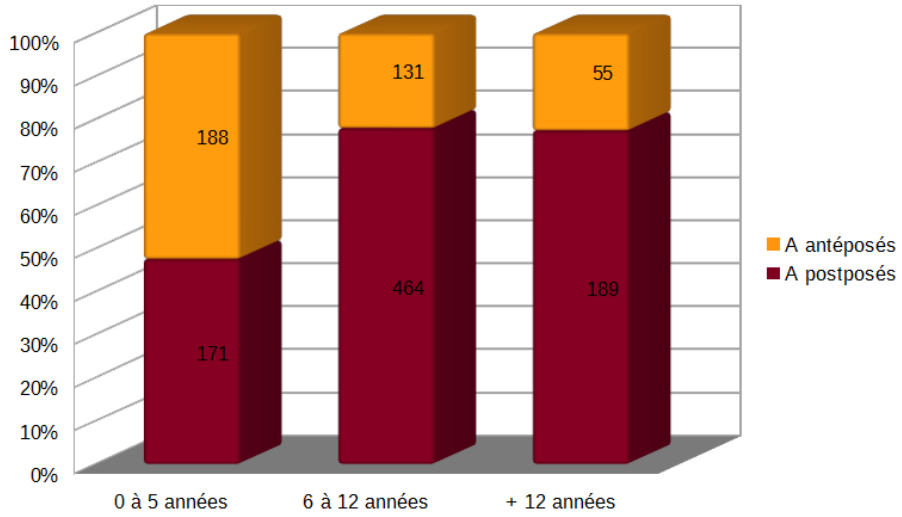


Figure 58 : Proportion d'adjectifs anté- et postposés produits par les participants des tests 1 et 2 en fonction de l'EDUCATION

La postposition de l'adjectif est pratiquée en très forte proportion, plus de 77 %, par les locuteurs ayant suivi plus de 5 années de scolarité, contrairement à ceux les moins scolarisés, qui la pratiquent à seulement 47,63 %.

3.4.2 Témoins

3.4.2.1 Structure des syntagmes nominaux définis

La distribution des syntagmes nominaux définis incluant un adjectif épithète en fonction de l'EDUCATION est représentée en Figure 59 (cf. Tableau 140).

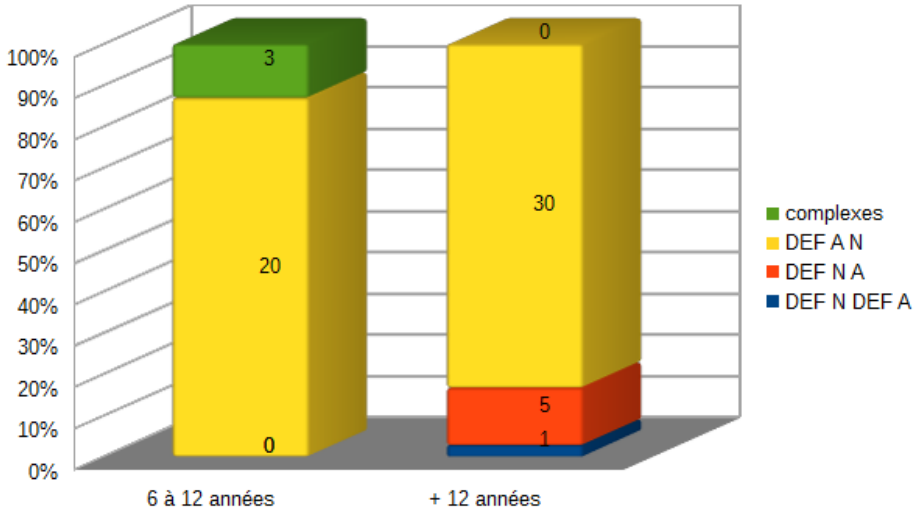


Figure 59 : Structure des SN définis produits par les témoins en fonction de l'EDUCATION

Les témoins ont un niveau d'éducation moyen et supérieur. La personne de la classe n°2 se comporte comme l'une des deux personnes de la classe n°3, c'est-à-dire qu'elle ne produit pas de syntagme DEF N DEF A, tandis qu'un DEF N DEF A est produit par une personne de la classe n°3 mais pas par l'autre. Tous trois produisent une grande majorité de syntagmes monodéfinis à adjectif antéposé : on constate que les syntagmes monodéfinis avec adjectif postposé sont produits seulement par les locuteurs ayant fait au moins 12 années de formation scolaire (2 occurrences chacun). Les deux personnes les plus scolarisées sont les seules à postposer l'adjectif (DEF N A ou DEF N DEF A).

3.4.2.2 Proportion d'adjectifs anté- et postposés

La distribution des syntagmes nominaux définis présentant un adjectif épithète anté- ou postposé en fonction de l'EDUCATION est représentée en Figure 60 (cf. Tableau 141).

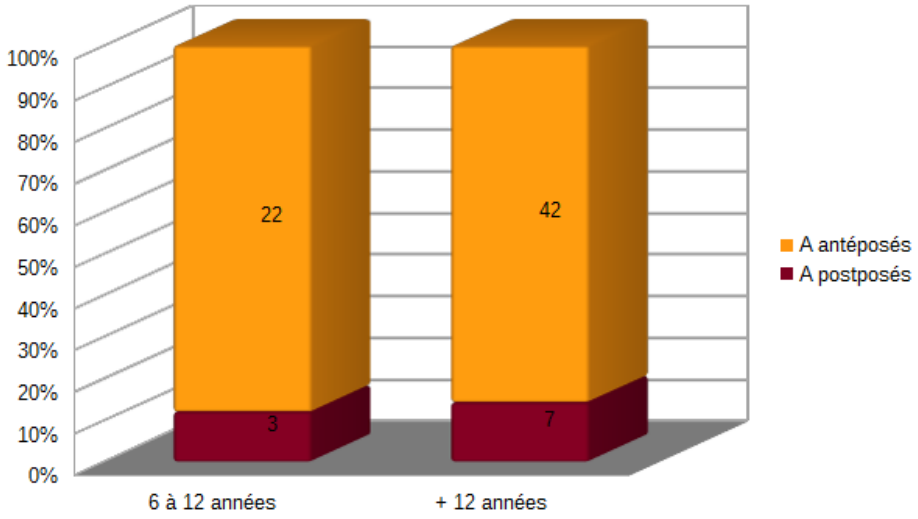


Figure 60 : Proportion d'adjectifs anté- et postposés produits par les témoins en fonction de l'EDUCATION

La postposition de l'adjectif n'est quasiment pas pratiquée par les témoins : 3 occurrences chacun, soit autour de 12 % chacun, quel que soit le niveau d'éducation. La personne la moins instruite postpose à elle seule autant ses adjectifs que les deux plus éduquées qu'elle.

On ne peut tirer de généralité en ce qui concerne la part des syntagmes polydéfinis et des syntagmes monodéfinis : les premiers sont produits à la fois par une personne du niveau d'éducation n°2 et par une personne du niveau d'éducation n°3, tandis qu'une personne du niveau d'éducation n°3 n'en produit pas du tout. On peut noter que les deux témoins plus instruits ont un comportement qui diffère légèrement de celui moyennement éduqué (niveau n°2), mais il est difficile de savoir si la variable EDUCATION suffit à l'expliquer.

3.5 Variable 4 : MILITANTISME

3.5.1 Participants albanais

3.5.1.1 Structure des syntagmes nominaux définis

La distribution des syntagmes nominaux définis incluant un adjectif épithète en fonction du MILITANTISME est représentée en Figure 61 (cf. Tableau 142).

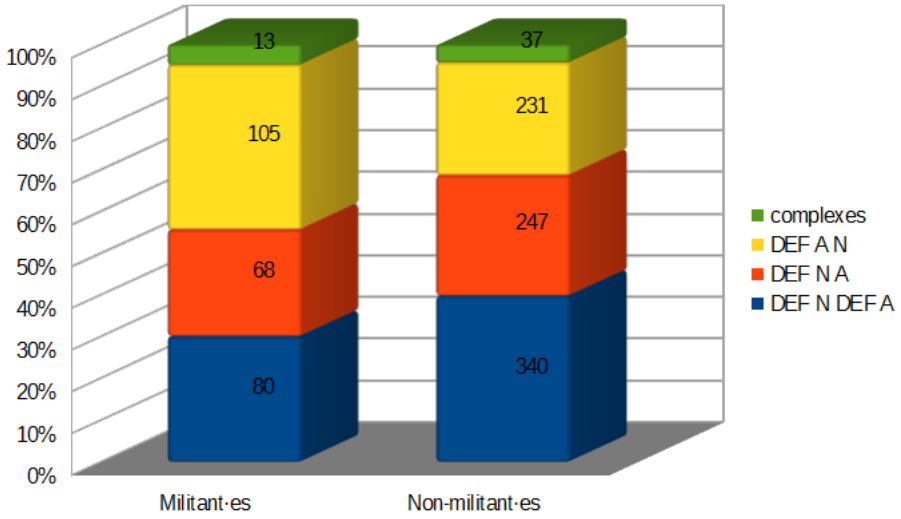


Figure 61 : Structure des SN définis produits par les participants des tests 1 et 2 en fonction du MILITANTISME

On observe que les non-militants produisent plus de syntagmes polydéfinis que les militants – 340 occurrences et 80, soit 39,76 % et 30,08 % de leur production respective. Les militants préfèrent antéposer l'adjectif dans 39,47 % des syntagmes monodéfinis contre 25,56 % de postposition (en syntagme monodéfini). Les non-militants postposent légèrement plus (29,89 %) qu'ils n'antéposent (27,02 %). L'emploi des trois structures est finalement bien réparti chez les deux profils, avec une impression de trois « tiers » assez équilibrés. Une structure a la préférence pour les deux profils, avec un pourcentage d'emploi similaire : DEF A N chez les militants (39,47 %) et DEF N DEF A chez les non-militants (39,76 %).

3.5.1.2 Proportion d'adjectifs anté- et postposés

La distribution des syntagmes nominaux définis présentant un adjectif épithète anté- ou postposé en fonction du MILITANTISME est représentée en Figure 62 (cf. Tableau 143).

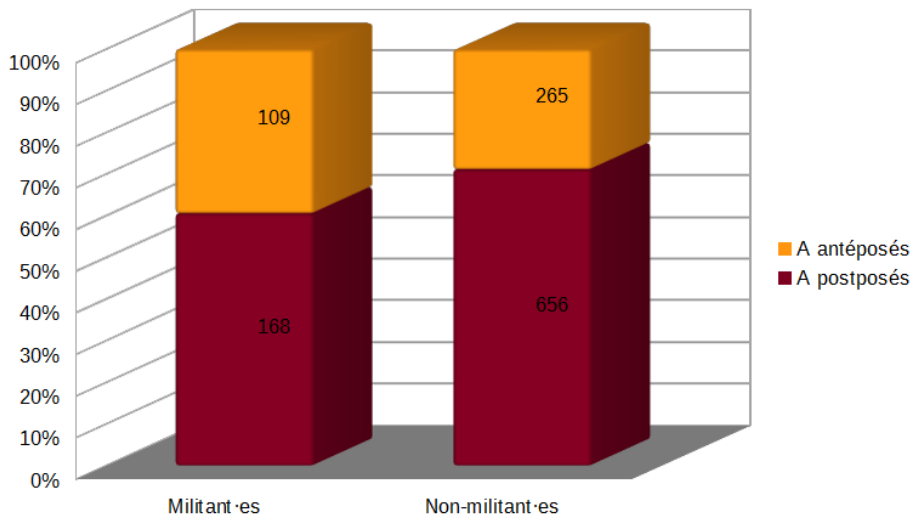


Figure 62 : Proportion d'adjectifs anté- et postposés produits par les participants des tests 1 et 2 en fonction du MILITANTISME

Il existe une différence entre les militants et les non-militants en ce qui concerne la place globale de l'adjectif dans le syntagme nominal : si l'ensemble des personnes postpose l'adjectif dans la majorité des cas, il s'agit d'un phénomène plus fort chez les non-militants (71,15 %) plutôt que chez les militants (60,65 %).

3.5.2 Témoins

3.5.2.1 Structure des syntagmes nominaux définis

La distribution des syntagmes nominaux définis incluant un adjectif épithète en fonction du MILITANTISME est représentée en Figure 63 (cf. Tableau 144).

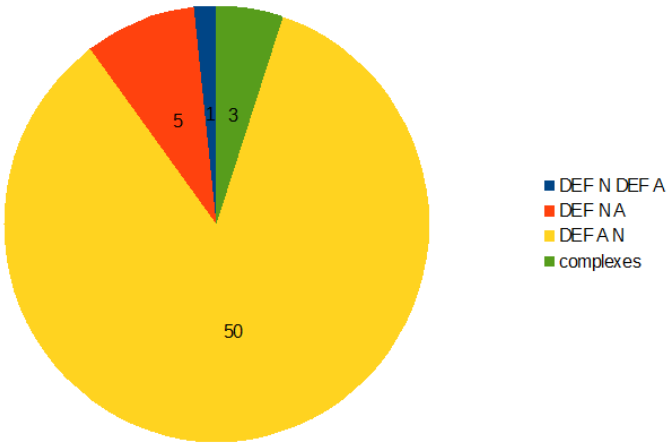


Figure 63 : Structure des SN définis produits par les témoins en fonction du MILITANTISME

Les trois personnes témoins ne sont pas militantes. On peut donc considérer leurs données comme un seul bloc. Contrairement aux participants albanais des deux tests, ces témoins non-militants ne produisent quasiment aucune structure DEF N DEF A. Or, contrairement aux participants non-militants albanais, ils antéposent en très grande majorité l'adjectif au sein des syntagmes monodéfinis, à 86,44 % contre 6,78 % de postposition.

3.5.2.2 Proportion d'adjectifs anté- et postposés

La distribution des syntagmes nominaux définis présentant un adjectif épithète anté- ou postposé en fonction du MILITANTISME est représentée en Figure 64 (cf. Tableau 145).

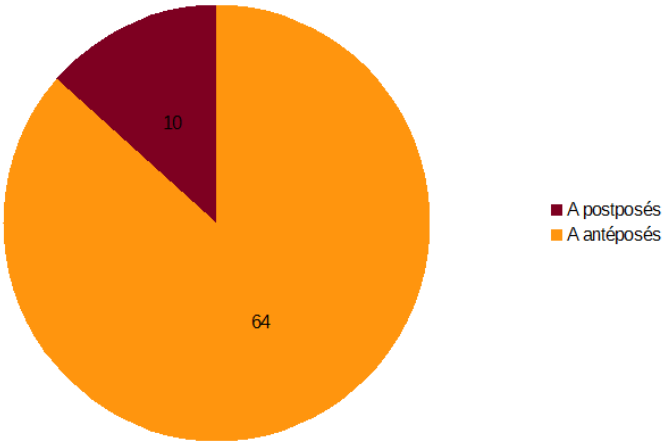


Figure 64 : Proportion d'adjectifs anté- et postposés produits par les témoins en fonction du MILITANTISME

Ils produisent chacun 3 adjectifs postposés avec d'autres déterminants, soit environ 11 à 13 % du total. Tous les autres syntagmes présentent un adjectif antéposé. La part des adjectifs antéposés de leur corpus représente donc 87,67 %, ce qui représente le contraire du résultat des non-militants.

3.6 Variable 5 : VARIETE

3.6.1 Participants albanais

3.6.1.1 Structure des syntagmes nominaux définis

La distribution des syntagmes nominaux définis incluant un adjectif épithète en fonction de la VARIETE est représentée en Figure 65 (cf. Tableau 146).

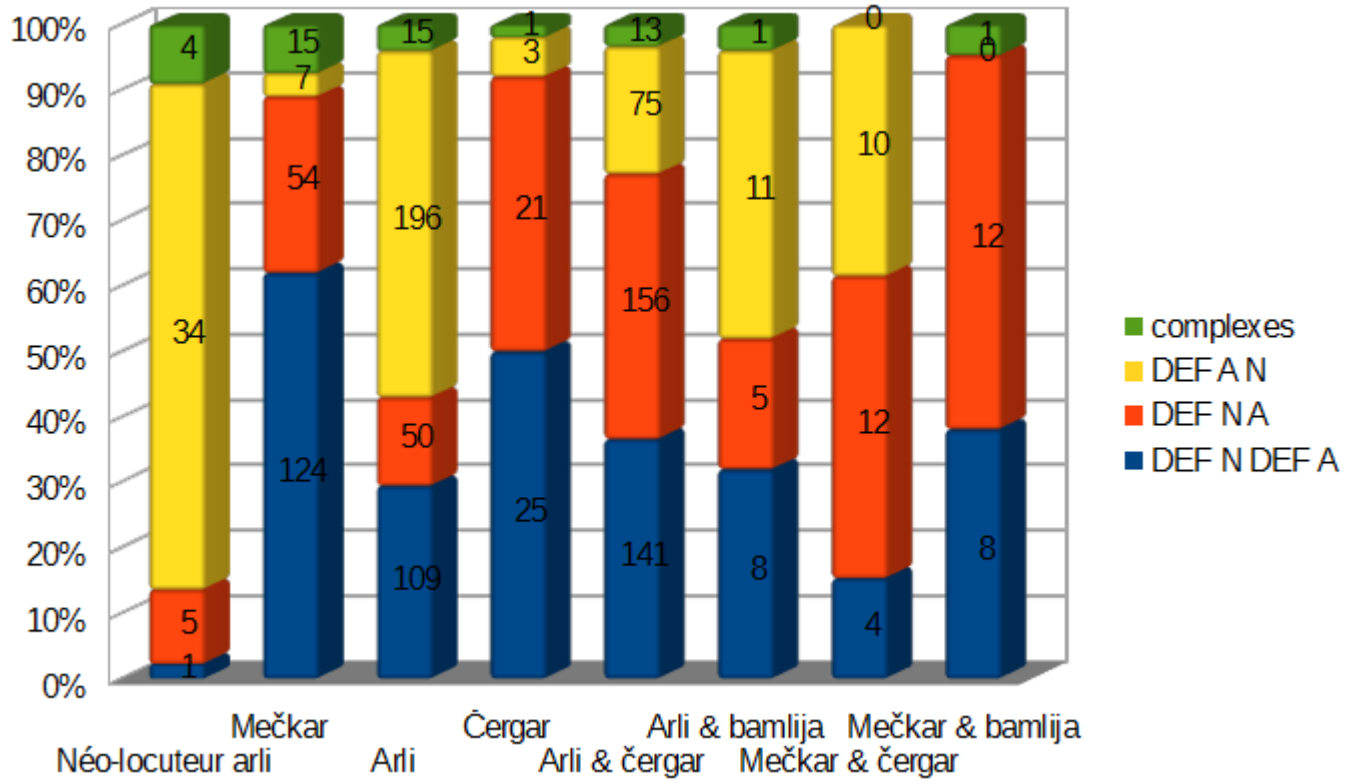


Figure 65 : Structure des SN définis produits par les participants des tests 1 et 2 en fonction de la VARIETE

Les locuteurs de la variété mečkar ont produit le plus de structures DEF N DEF A (62 %), soit la majorité des syntagmes nominaux définis. Viennent ensuite les locuteurs de la variété čergar I (50 %) et arli (29,46 %). Quant aux personnes bivariétales, on constate que Mečkar & bamlija présente un taux relativement haut, le plus haut des bivariétales (38,10 %), puis suivent les Arlis & čergars I, situés entre la moyenne des Arlis et celle des Čergars I (36,62 %). La personne Arli & bamlija se trouve quant à elle proche des Arlis (32 %). De façon étonnante, la personne Mečkar & čergar I présente un taux très bas, bien loin de la moyenne des Mečkars comme de celle des Čergars I (15,38 %). Le néo-locuteur arli ne produit qu'une fois la structure (2,27 %).

En ce qui concerne la place de l'adjectif dans les syntagmes monodéfinis, la postposition est pratiquée par les Čergars I (42 %) et les Arlis & čergars I (40,52 %) de façon massive, ainsi que par les Mečkars (27 %). De façon étonnante, les personnes bivariétales ont tendance à plus postposer l'adjectif que celles monovariétales : la personne Mečkar & bamlija plus (57,14 %) que les Mečkars à une variété (27 %), la personne Mečkar & čergar I plus (46,15 %) que les Mečkars à une variété (27 %) et que les Čergars I à une variété (42 %). Les locuteurs qui postposent plus qu'ils ne antéposent l'adjectif (DEF N DEF A + DEF N A ~ DEF A N) sont donc les Mečkars et les Čergars I, ainsi que les personnes monovariétales parlant Mečkar & bamlija, Arli & čergar I ou Mečkar & čergar I. Cela explique que, dans le test 1, nous trouvions également ces résultats pour les Mečkars et les Arlis & čergars I. Les locuteurs des autres variétés, arli et comprenant l'arli, antéposent au contraire en majorité l'adjectif : à 52,97 % pour les Arlis à une variété (contre 13,51 % de postposition) et à 77,27 % pour le néo-locuteur de l'arli (contre 11,37 % de postposition), à 44 % pour le locuteur Arli & bamlija (contre 20 % de postposition).

3.6.1.2 Proportion d'adjectifs anté- et postposés

La distribution des syntagmes nominaux définis présentant un adjectif épithète anté- ou postposé en fonction de la VARIETE est représentée en Figure 66 (cf. Tableau 147).

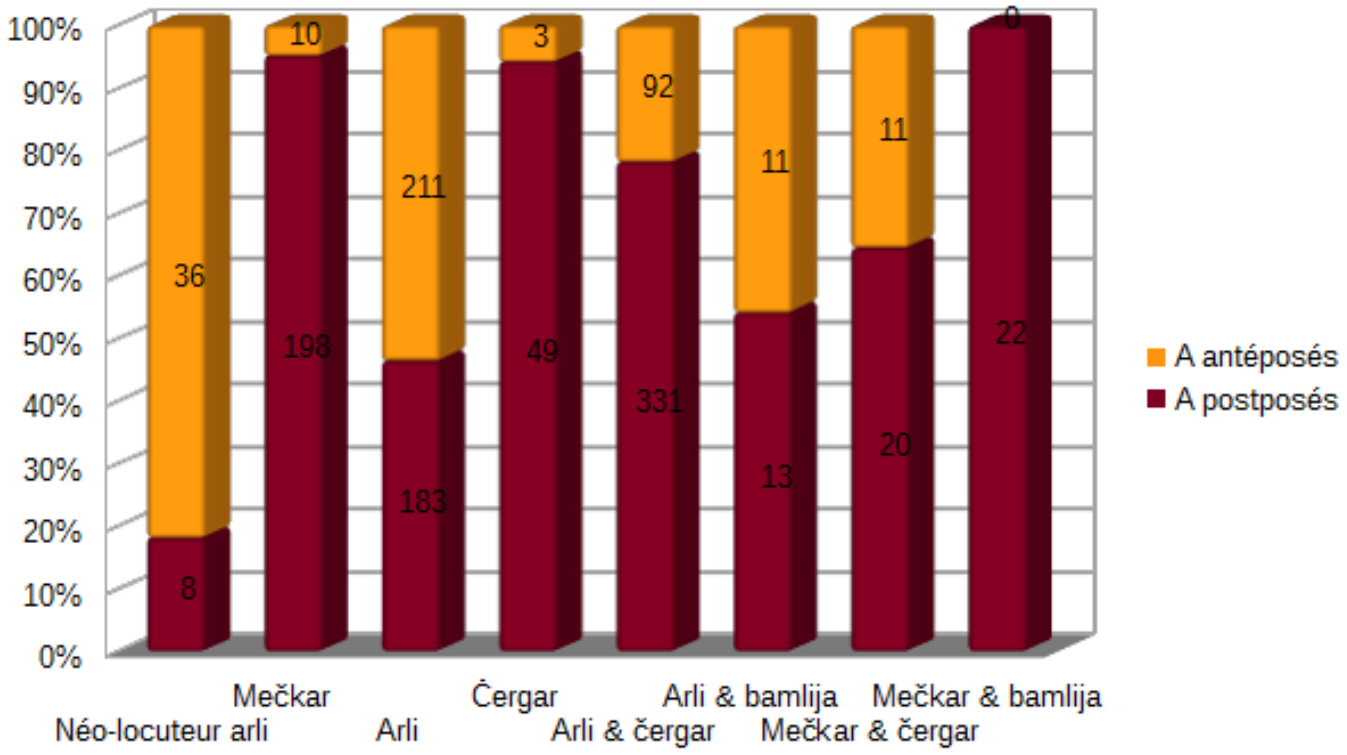


Figure 66 : Proportion d'adjectifs anté- et postposés produits par les participants des tests 1 et 2 en fonction de la VARIETE

La postposition de l'adjectif est pratiquée de manière assez différenciée en fonction des variétés. Le néo-locuteur arli est le seul à antéposer en grande majorité ses adjectifs (81,82 %), tandis que les Arlis proprement dit antéposent en courte majorité l'adjectif (53,55 %). Proche des Arlis se trouve le locuteur bi-variétal Arli & bamlija qui antépose à 45,83 %, contre 54,17 % de postposition. Les autres catégories présentent un taux très élevé et majoritaire de postposition de l'adjectif : Arlis & čergars I (78,25 %), Čergars I (94,23 %) et Mečkars (95,19 %). Les Arlis & čergars I sont donc bien situés entre les Arlis et les Čergars I à une variété. Le Mečkar & bamlija a quant à lui quasiment le même taux que les Mečkars à une variété (100 %). Là aussi, le locuteur mečkar & čergar I se trouve loin des Mečkars et des Čergars I à une variété, avec seulement 64,52 %.

3.6.2 Témoins

3.6.2.1 Structure des syntagmes nominaux définis

La distribution des syntagmes nominaux définis incluant un adjectif épithète en fonction de la VARIETE est représentée en Figure 67 (cf. Tableau 148).

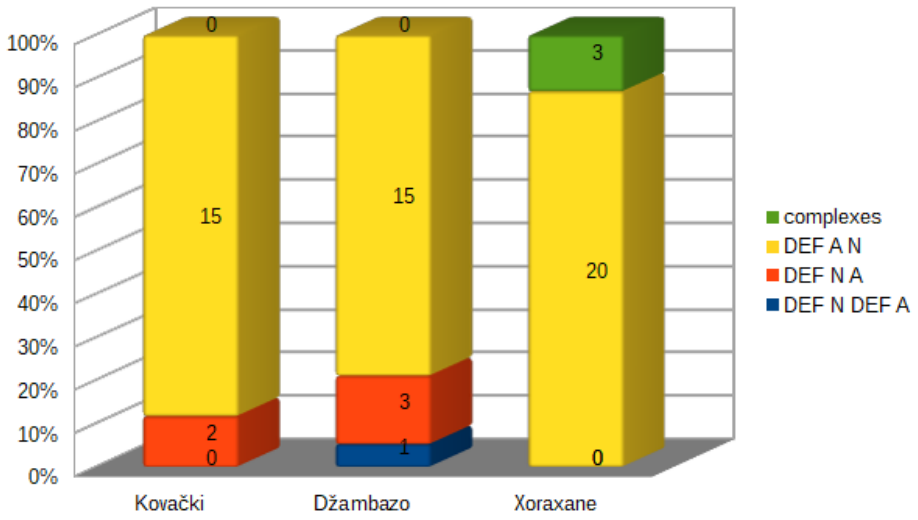


Figure 67 : Structure des SN définis produits par les témoins en fonction de la VARIETE

Les témoins parlent trois variétés différentes, l'une du Kosovo (xoraxane) et deux de Macédoine (kovački et džambazo). Un seul, originaire de Macédoine, produit un syntagme DEF N DEF A. Les deux Macédoniens ont en commun de produire quelques DEF N A (à hauteur d'un peu plus de 10 %). Enfin, tous ont en commun d'antéposer en très grande majorité l'adjectif monodéfini.

3.6.2.2 Proportion d'adjectifs anté- et postposés

La distribution des syntagmes nominaux définis présentant un adjectif épithète anté- ou postposé en fonction de la VARIETE est représentée en Figure 68 (cf. Tableau 149).

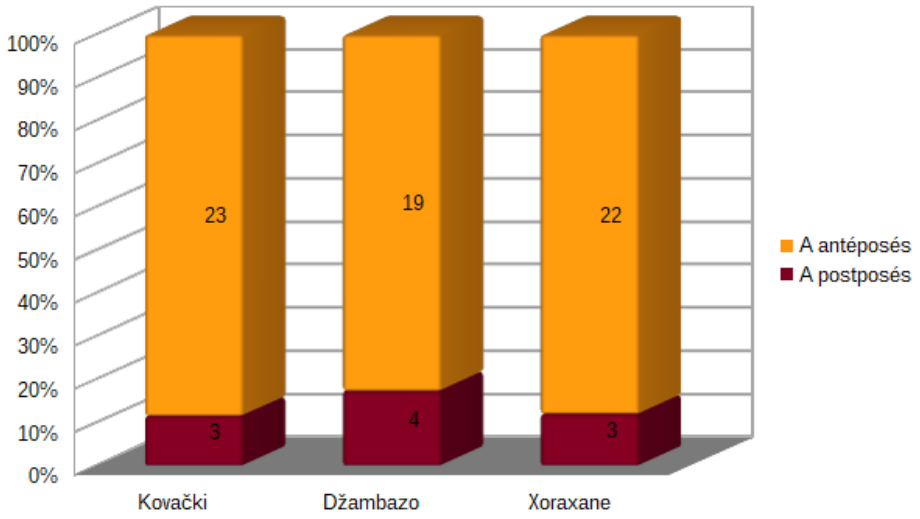


Figure 68 : Proportion d'adjectifs anté- et postposés produits par les témoins en fonction de la VARIETE

Les témoins, toutes variétés confondues, produisent chacun 3 adjectifs postposés avec d'autres déterminants, soit environ 11 à 13 % – il ne semble pas ici que les variétés diffèrent fortement entre elles.

3.7 Synthèse des variables sociales

3.7.1 Test 1 ~ test 2

Pour des raisons de place, nous avons présenté les résultats des deux versions du test de manière unitaire. Existe-t-il une différence importante entre les résultats du test 1 et ceux du test 2 ?

- GENRE : on observe une différence. Les femmes produisent nettement plus de DEF N DEF A que les hommes dans le test 1, et cet écart est moindre dans le test 2. Elles produisent nettement moins de DEF N A que les hommes dans le test 1, et cet écart est moindre dans le test 2. On ne note pas de différence sensible quant à la place de l'adjectif quel que soit le déterminant.
- CLASSE D'AGE : on observe peu de différences, à l'exception des plus de 60 ans. Ils sont représentés par un seul locuteur dans le test 1, qui produit beaucoup plus de DEF N DEF A que ses homologues dans le test 2. Il produit aussi plus de DEF N A mais beaucoup moins de DEF A

N. La différence est nette aussi quant à la place de l'adjectif, beaucoup plus postposé chez lui que chez ses homologues du test 2.

- EDUCATION : on observe peu de différences, à l'exception des peu/non scolarisés (0-5 années). Ils présentent un fort taux de présence de DEF N DEF A dans le test 1 par rapport au test 2. Ils produisent plus de DEF N A dans le test 1 que dans le test 2, et beaucoup moins de DEF A N dans le test 1 que dans le test 2. La différence est très nette également en ce qui concerne la place de l'adjectif : beaucoup plus de postposition de l'adjectif dans le test 1 que dans le test 2.
- MILITANTISME : on observe une différence. L'écart entre militants et non-militants est plus marqué pour DEF N DEF A dans le test 1 que dans le test 2. Si, dans le test 1, les militants produisent moins de DEF N A que les non-militants, les militants en produisent autant que les non-militants dans le test 2. Si, dans le test 1, les militants produisent seulement beaucoup plus de DEF A N que les non-militants, les militants en produisent un peu plus que les non-militants dans le test 2. L'adjectif est nettement moins postposé par les militants dans le test 1 que dans le test 2.
- VARIETE : on observe une différence. Le néo-locuteur arli ne produit pas du tout de DEF N DEF A dans le test 1 et 1 seul dans le test 2. Il produit nettement plus de DEF N A dans le test 1 que dans le test 2. Les Mečkars produisent proportionnellement plus de DEF N DEF A dans le test 1 que dans le test 2. Les Arlis produisent proportionnellement plus de DEF N DEF A dans le test 1 que dans le test 2 et moins de DEF A N dans le test 1 que dans le test 2. Enfin, le néo-locuteur arli pratique plus la postposition de l'adjectif dans le test 1 que dans le test 2. Les Arlis font le contraire dans le test 1 de ce qu'ils font dans le test 2 : la postposition de l'adjectif majoritaire dans le test 1 et minoritaire dans le test 2.

On voit que le test 1 présente une différence importante pour les modalités suivantes : femmes, plus de 60 ans, personnes peu/non scolarisées, militants, Arlis. On peut considérer qu'il n'y a pas hétérogénéité entre les deux versions du test, mais une différence de degré : la variation est plus intense dans le test 1 que dans le test 2. À l'exception du locuteur de plus de 60 ans du test 1, qui est le seul représentant de sa modalité, et à l'exception des militants, dont la pertinence peut être remise en cause dans le test 2.

3.7.2 Hiérarchie des variables

Les variables ne sont pas toutes des facteurs d'explication. Les Figure 61 et Figure 62 ne dégagent pas de groupes ayant un comportement syntaxique différent : cela montre que la variable MILITANTISME n'est pas très pertinente ou bien que le découpage effectué n'était pas pertinent. Les autres variables n'ont pas non plus toutes le même poids :

- EDUCATION produit des résultats différents dans les tests 1 et 2 : les modalités sont hiérarchisées de manière différente dans les tests 1 et 2, c'est-à-dire que les groupes de locuteurs ont un comportement syntaxique parfois différent d'un test à l'autre.
- GENRE produit des différences significatives dans le test 1 mais moins dans le test 2.

- CLASSE D'AGE produits des résultats similaires dans les tests 1 et 2.
- VARIETE produits des résultats similaires dans les tests 1 et 2 – même si la lecture des résultats est rendue plus difficile en raison du grand nombre de modalités qui la composent.

On peut classer les variables sociales par ordre décroissant d'impact (Figure 69) :

VARIETE - CLASSE D'AGE > GENRE > EDUCATION > MILITANTISME

Figure 69 : Ordre d'impact décroissant des variables sociales

3.7.3 Profil des locuteurs

3.7.3.1 Participants albanais

Quel locuteur est le plus susceptible de produire quelle structure ? C'est l'objet de la Figure 70.

Tableau 93 : Profil de locuteur le plus probable pour chaque syntaxe nominale

	GENRE	CLASSE D'AGE	EDUCATION	MILITANTISME	VARIETE
DEF N DEF A	femme	15-59 ans	0-5 années $_{test\ 1}$/ +12 années $_{test\ 2}$	non-militant	mečkar, čergar I
DEF N A	homme	- 15 ans	6-12 années	non-militant	čergar bi-variétal, mečkar
DEF A N	=	+ 60 ans	6-12 années $_{test\ 1}$ / 0-5 années $_{test\ 2}$	militant	arli & arli bivariétal
adjectif postposé	=	- 15 ans	0-5 années $_{test\ 1}$ / 6-12 années $_{test\ 2}$	non-militant	mečkar & mečkar bivariétal
adjectif antéposé	=	+ 40 ans	6-12 années $_{test\ 1}$ / 0-5 années $_{test\ 2}$	militant	arli & arli bivariétal

La personne qui a le plus de probabilité de produire fréquemment DEF N DEF A est donc une femme entre 15 et 59 ans, non-militante, parlant la variété mečkar ou čergar I. Celle qui a le plus de probabilité de produire fréquemment DEF N A est un homme de moins de 15 ans, moyennement scolarisé, non-militant, parlant čergar I, mečkar & čergar I, arli & čergar I ou mečkar. Celle qui a le plus de probabilité de produire fréquemment DEF A N a plus de 60 ans, est militant et parle arli ou néo-locuteur arli, arli & bamlija, arli & čergar I. Ce peut également être l'un des témoins. La personne qui a le plus de probabilité de postposer l'adjectif a moins de 15 ans, ne milite pas, parle mečkar ou une variété comprenant mečkar ou mečkar & čergar I, mečkar & bamlija. Celle qui a le plus de probabilité d'antéposer l'adjectif a plus de 40 ans, milite et parle arli ou néo-locuteur arli, arli & bamlija, arli & čergar I.

3.7.3.2 Témoins

Les témoins présentent peu de différences entre eux, quelle que soit la variable considérée. S'il est vrai qu'il ne s'agit que de trois personnes, leurs résultats sont très semblables entre eux, ce qui laisse à penser qu'il existe bel et bien une syntaxe nominale des variétés albanaises distinctes de celles, par exemple, de Macédoine et du Kosovo. Deux témoins ont en commun de postposer l'adjectif plus que le troisième : comme ils appartiennent à la même CLASSE D'AGE, au même niveau d'EDUCATION et parlent deux VARIETES en contact avec la même langue dominante (macédonien), difficile d'établir laquelle de ces variables explique leur similitude.

Si les témoins présentent une homogénéité entre eux, ils présentent une forte hétérogénéité avec l'ensemble des locuteurs d'Albanie. Globalement, ils ne pratiquent quasiment pas la polydéfinitude ni la postposition de l'adjectif, avec article défini ou non. Les locuteurs auxquels ils ressemblent sont :

- les moins de 15 ans et les plus de 60 ans en ce qui concerne la forte tendance à la monodéfinitude
- les plus de 60 ans en ce qui concerne la forte tendance à antéposer l'adjectif
- les Arlis et le néo-locuteur arli en ce qui concerne la forte tendance à la monodéfinitude et à l'antéposition de l'adjectif

Les témoins se différencient des autres par rapport à la VARIETE et à la CLASSE D'AGE.

3.7.3.3 Regroupements de profils

On peut recouper les modalités de différentes variables pour affiner les profils dégagés. On remarque que les témoins ont une syntaxe proche des plus de 60 ans et du néo-locuteur arli : peu ou pas de DEF N DEF A et de DEF N A, et un DEF A N quasiment catégorique. Ils sont proches également des moins de 15 ans en ce qui concerne la faible présence de DEF N DEF A et la forte tendance à la monodéfinitude, même si bien sûr la place de l'adjectif est inverse chez ces catégories de locuteurs. Chez les Mečkars, les Čergars I et les Arlis & čergars I, la stratégie adoptée est en grande partie DEF N DEF A et en moindre partie DEF N A – avec une proportion de DEF N DEF A qui ressemble plus à ce que l'on trouve chez les classes d'âge moyennes, chez les femmes, les personnes moins scolarisées et chez les Arlis. Ces regroupements sont représentés en Figure 70 (cf. Tableau 150).

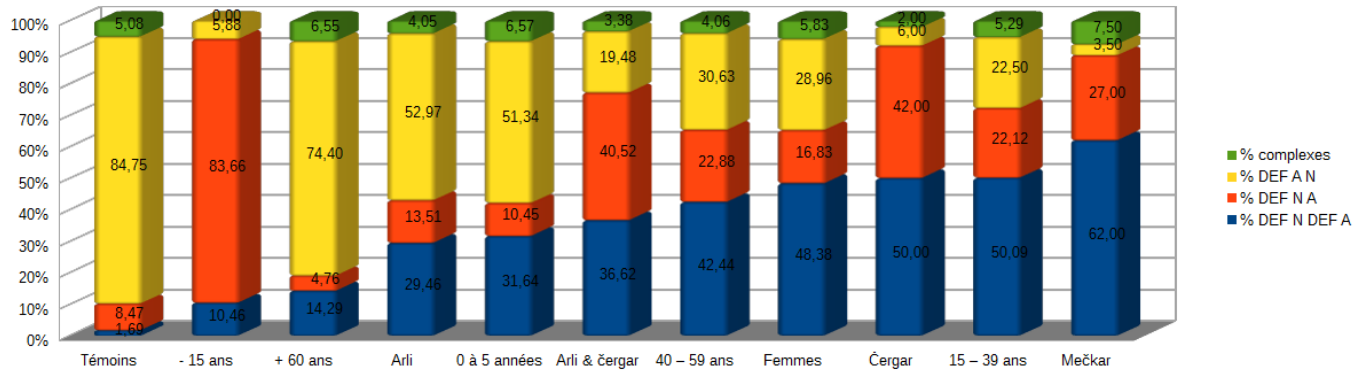


Figure 70 : Regroupement des modalités convergentes issues de GENRE, CLASSE D'AGE, EDUCATION et VARIETE dans le pourcentage de production des différents SN définis

En ce qui concerne la place de l'adjectif quel que soit le déterminant, les témoins ressemblent aux plus de 60 ans ; ils s'opposent totalement aux moins de 15 ans, aux femmes, aux personnes moins scolarisées et aux autres variétés. Ces regroupements sont représentés en Figure 71 (cf. Tableau 151).

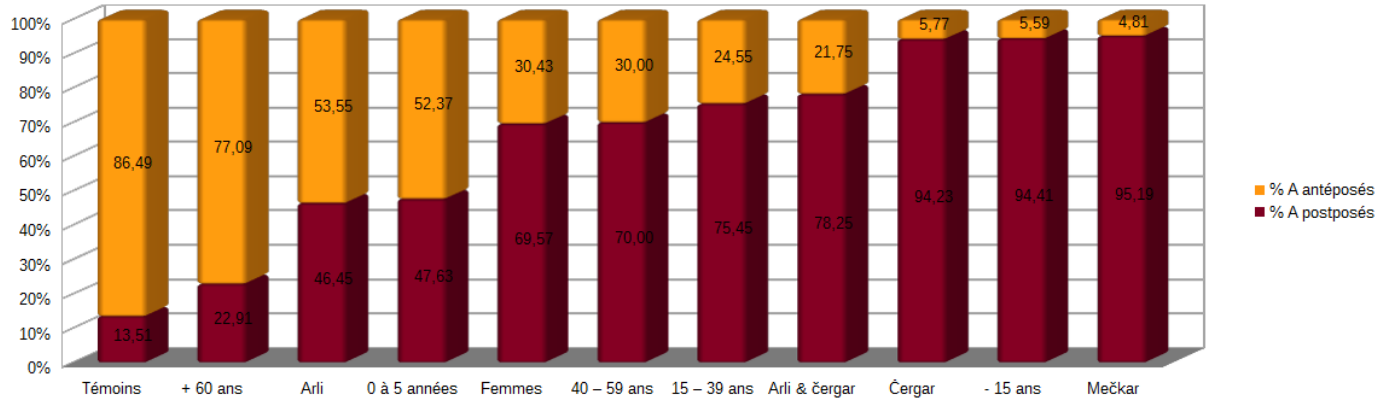


Figure 71 : Regroupement des modalités convergentes issues de GENRE, CLASSE D'AGE, EDUCATION et VARIETE dans le pourcentage de production d'adjectifs anté- et postposés

3.7.4 Perspectives

3.7.4.1 Difficultés

On peut émettre deux objections quant aux résultats obtenus :

1) Nous avons considéré les variables de façon indépendante, alors que l'explication peut venir d'une combinaison de variables. Si, par exemple, une structure n'était produite que par les jeunes hommes de moins de 30 ans, une explication isolée par CLASSE D'AGE ne produirait pas de résultat

concluant, car la classe d'âge des moins de 30 ans, par exemple, mêlerait des locuteurs produisant le phénomène à des locutrices ne le produisant pas. Une étude sur l'ensemble des locuteurs du corpus en fonction de GENRE, de même, mélangerait les hommes de moins de 30 ans qui produisent la structure à ceux de plus de 30 ans qui ne la produisent pas.

2) Nous avons considéré jusqu'ici les modalités des variables en bloc, sans tenir compte des individus qui la composent. Il est toujours possible que certains locuteurs ne soient pas représentatifs de leur catégorie – ce qui peut s'envisager pour le locuteur mečkar & čergar I, dont les résultats sont très loin à la fois de ceux des Mečkars et de ceux des Čergars I. Tagliamonte et Baayen (2012 : 9) parlent d'*atypical outliers* :

> « *There are strategies and even a rich literature on the anomalous behaviors of individuals inside community-based samples (e.g. lames (Labov, 1972b), oddballs (Chambers, 1998, p. 94)), and strategies have been proposed to find and evaluate the effect such individuals may have on the data (van de Velde and van Hout, 1998; Guy, 1980). Although mixed-effects models can bring individual differences into the statistical model, they do not protect protection against distortion by atypical outliers. Model criticism is an essential part of good statistical practice, irrespective of whether a mixed-effects approach is adopted.* » (Tagliamonte et Baayen 2012 : 9)

Certains individus peuvent poser problème, s'ils ne pratiquent pas la variation mais choisissent toujours la même option : on parle d'individus « catégoriques » par rapport aux individus « variables ». Certaines études préfèrent les écarter du corpus, afin d'obtenir des groupes homogènes, et donc des résultats plus cohérents – une moyenne comprenant la majorité homogène des locuteurs et ces *outliers* et individus catégoriques n'aurait pas de sens, ou une valeur seulement abstraite. Tagliamonte et Baayen (2012 : 27) soulèvent le problème :

> « *Variationist methodology typically recommends that categorical individuals be removed for the study of variable phenomena (e.g., Guy, 1988, p. 130). However, in practice, particularly with morpho-syntactic and discourse-pragmatic features, they are often included on the assumption that internal predictors will be parallel across individuals. The question is whether or not these individuals without variation are a source of noise that should be taken out before the start of the analysis? Would the relative importance of the predictors change if a random forest were fitted to the data after exclusion of the non-variable individuals?* » (Tagliamonte et Baayen 2012 : 27)

3.7.4.2 Solutions

Nous proposons quelques pistes de réflexions pour surmonter les difficultés évoquées :

1) Le premier problème concerne les effets mixtes. Si VARIETE et CLASSE D'AGE sont les deux facteurs les plus explicatifs, il faut les combiner dans les calculs et pour cela, établir un échantillon de locuteurs stratifié en fonction de ces deux variables particulières, où chaque modalité d'âge comprend un nombre égal de locuteurs de chaque variété, et où chaque variété comprend autant de locuteurs de chaque tranche d'âge. Pour des résultats plus nets, il conviendrait d'exclure les locuteurs « à deux variétés ». Concrètement, cela suppose un séjour plus long sur le terrain ou bien la décision de salarier les locuteurs pour leur participation au test, afin de pouvoir disposer de suffisamment de

locuteurs et de pouvoir les sélectionner en fonction de leur profil. On adoptera ensuite un modèle statistique qui tienne compte de l'intersection des variables et puisse les croiser toutes. Il existe trois types principaux de modèles en statistique : le modèle linéaire généralisé (*generalized linear model*), le modèle linéaire généralisé à effets mixtes (*generalized linear mixed-effects model*), et enfin le modèle des forêts aléatoires (*random forest*) utilisant des arbres d'inférence conditionnelle (*conditional inference trees*) (Tagliamonte et Baayen 2012 : 32) :

> « *In general, the mixed-effects model is an excellent choice for relatively balanced data sets with one or more, potentially crossed, random effect factors (individuals, words, constituents, etc.). For highly unbalanced designs and complex interactions, conditional inference trees and random forests are more flexible, and may yield superior models. However, for large data sets with multiple random-effect factors with many levels, they rapidly become computationally intractable, given current hardware.* » (Tagliamonte et Baayen 2012 : 32)

2) Pour résoudre la difficulté posée par les *outliers* et par les individus catégoriques, il convient de représenter les individus au sein des modalités, afin de mettre en évidence ceux dont les résultats sont très éloignés des autres. Il convient également de représenter l'ensemble du corpus par individu et non par variable, afin de voir si des individus se distinguent du reste de leur groupe. Cela permettra également de confirmer ou d'infirmer la validité des modalités choisies (notamment le découpage en niveaux d'éducation et la partition militants / non-militants). Enfin, on peut introduire une nouvelle variable: INDIVIDU, qui puisse prendre en compte les préférences idiosyncrasiques ou l'histoire personnelle de chacun des locuteurs. Tagliamonte et Baayen (2012 : 8) précisent pourquoi[1] :

> « *Consider the individual speakers or writers typically sampled in sociolinguistic studies. Such individuals are generally sampled from larger populations of similar individuals, and are selected to be representative for these larger populations (e.g., the city, region or dialect individuals come from). This brings us to an important distinction in the statistical analysis of factorial data (i.e. data that can be categorized into levels or factors). There is an important difference between fixed-effect and random-effect factorial predictors. An example of a fixed-effect factor is the sex of the individual, which has exactly two levels (female versus male) and exhausts all possible levels for this predictor, at least from a biological perspective. Random-effect predictors, by contrast, have levels that constitute only a subset of the possible categories available in the population. Individuals (and also words, e.g., nouns, verbs or adjectives) are typical examples of random-effect factors. If, in a statistical analysis, a random-effect predictor is analysed as if it were a fixed-effect predictor, then the conclusions reached will only be valid for the individuals and words sampled. Thus, if the sample comprises 8 individuals the statistical model will be valid for only those 8 individuals. Conclusions do not automatically generalize to the relevant populations of interest.* » (Tagliamonte et Baayen 2012 : 8)

INDIVIDU est la seule variable **à effet aléatoire** contrairement aux autres qui sont à effet fixe. Elle peut permettre d'éviter les erreurs, notamment sur un petit corpus collecté auprès d'un petit échantillon de personnes » (Tagliamonte et Baayen 2012 : 30) :

> « *Studies that do not bring Individual into the model specification not only run the risk of failing to come to grips with an important source of variation, they also run the risk of reporting a result as*

[1] *Predictor* est employé pour « variable », *level* pour « modalité ».

significant which upon closer inspection turns out not to be not significant, i.e. an anti-conservative interpretation of results […]. » (Tagliamonte et Baayen 2012 : 30)

3.7.4.3 Variable 6 : INDIVIDU

Structure des syntagmes nominaux définis

Nous avons observé les résultats individu par individu, en comparant les pourcentages respectifs de production de chacune des structures DEF A N, DEF N A et DEF N DEF A. Les résultats se trouvent en Figure 72 (cf. Tableau 152).

On trouve des individus catégoriques ou quasiment catégoriques des trois structures DEF N DEF A, DEF N A, DEF A N.

On constate que les trois témoins ont des résultats similaires à ceux du néo-locuteur arli M22a et des locuteurs arli M71A et F65A : peu ou pas de DEF N DEF A (respectivement 0 %, 0 % et 5,26 % pour les témoins M43X, F24K et M27Dž ; respectivement 0 %, 2,27 % et 2,60 % pour les Albanais M71A, M22a et F65A), peu ou pas de DEF N A (respectivement 0 %, 11,76 % et 15,79 % pour les témoins M43X, F24K et M27Dž ; respectivement 0 %, 11,36 % et 1,30 % pour les Albanais M71A, M22a et F65A), et un DEF A N quasiment catégorique (respectivement 86,96 %, 88,24 % et 78,95 % pour les témoins M43X, F24K et M27Dž ; respectivement 100 %, 77,27 % et 90,91 % pour les Albanais M71A, M22a et F65A).

Quatre personnes sont quasiment catégoriques dans leur production de DEF N A : M11A (arli, moins de 15 ans : 93,75 %), M19M(1) (mečkar, 15-39 ans : 86,67 %), M13AČ (arli & čergar I, moins de 15 ans : 83,58 %) et M58Č (čergar I, 40-59 ans : 86,36 %).

Quatre personnes sont quasiment catégoriques dans leur production de DEF N DEF A : M18Č (čergar I, 15-39 ans : 78,57 %), F20A (arli, 15-39 ans : 82,69 %), F53A (arli, 40-59 ans : 85,71 %), M45M (mečkar, 40-59 ans : 88 %) et F9A (arli, moins de 15 ans : 100%).

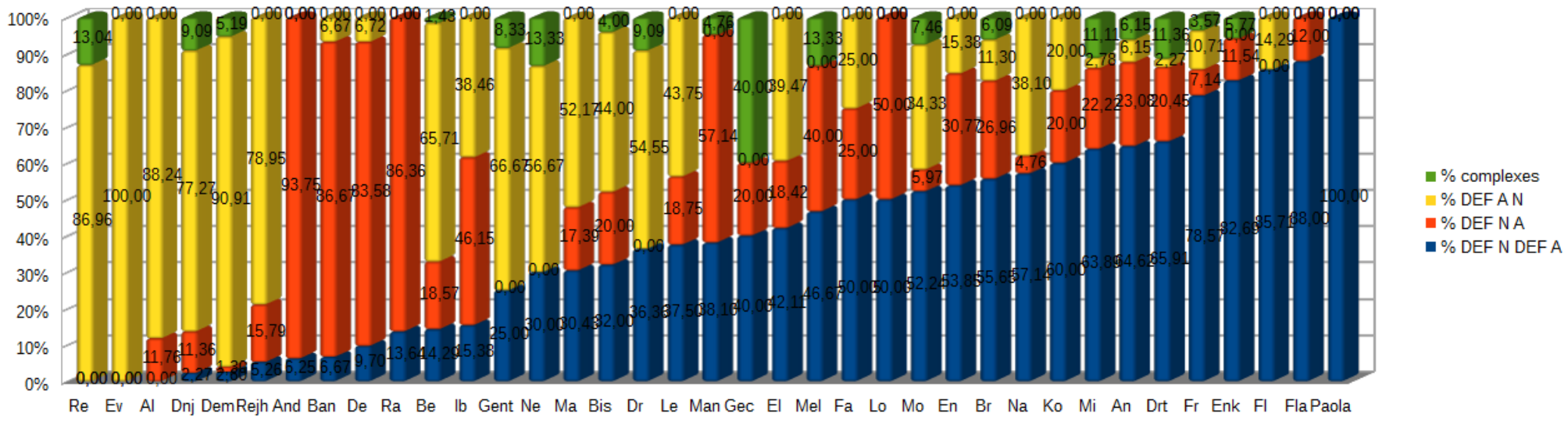


Figure 72 : Pourcentage de production des différents SN définis par les participants des tests 1 et 2 en fonction de INDIVIDU

Proportion d'adjectifs anté- et postposés

Nous avons observé les résultats individu par individu, en comparant les pourcentages respectifs de d'anté- et de postposition de l'adjectif. Les résultats se trouvent en Figure 73 (cf. Tableau 152).

On trouve les mêmes similarités pour la place de l'adjectif en général et pour les syntagmes définis. On constate que les trois témoins ont des résultats similaires à ceux du néo-locuteur arli M22a et des locuteurs arli M71A et F65A : peu ou pas d'adjectifs postposés (respectivement 11,54 %, 12 % et 17,39 % pour les témoins F24K, M43X et M27Dž ; respectivement 0 %, 6,49 % et 18,18 % pour les Albanais M71A, F65A et M22a), et un adjectif antéposé quasiment catégorique (respectivement 88,46 %, 88 % et 82,61 % pour les témoins F24K, M43X et M27Dž ; enfin respectivement 100 %, 93,51 % et 81,82 % pour les Albanais M71A, F65A et M22a).

Dix-neuf personnes sont quasiment catégoriques dans leur production d'adjectifs postposés, dans la mesure où elles le produisent à 80 % et plus : F22AČ, M52A, M37A, F16AČ, F53A, M18Č, M19M$_{(1)}$, M13AČ, M45M, F22M, F47M, M19M$_{(2)}$, M11A, F20A, F11A, F37MB, F66M, F9A, M58Č (dont 8 qui postposent l'adjectif dans 100% des cas). Parmi eux, on trouve les quatre personnes de moins de 15 ans, neuf personnes de 15 à 39 ans sur les seize, cinq personnes de 40 à 59 ans sur les huit, et une personne de plus de 60 ans sur les six. Parmi eux, on trouve les six personnes mečkar, la personne mečkar & bamlija, les deux personnes čergar I, sept personnes arli sur les quinze, trois personnes arli & čergar I sur les sept.

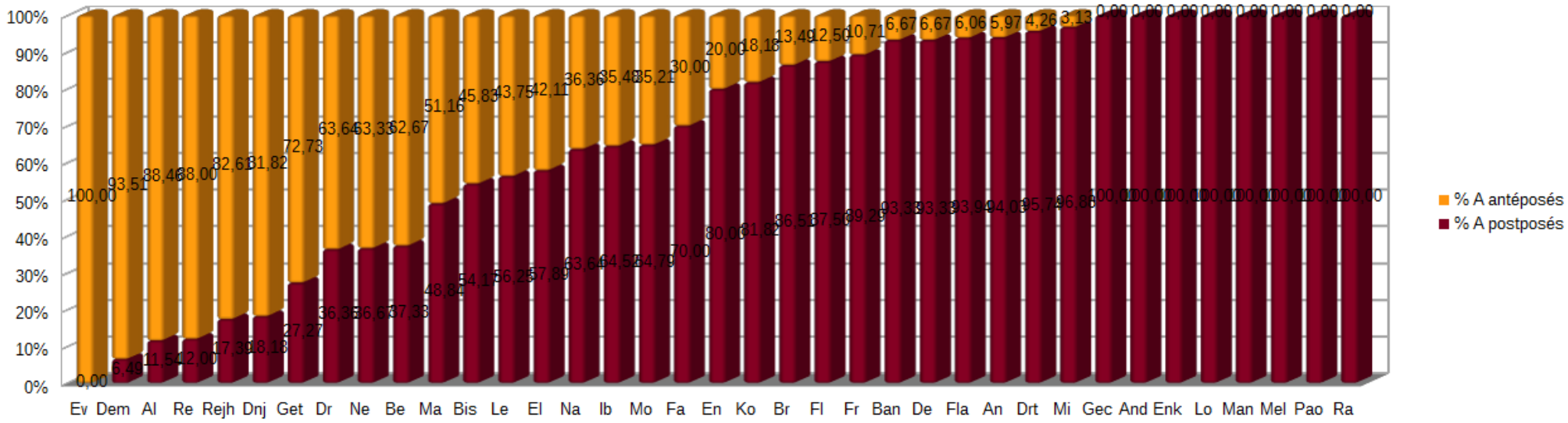


Figure 73 : Pourcentage de production d'adjectifs anté- et postposés par les participants des tests 1 et 2 en fonction de INDIVIDU

Bilan

Les témoins ont pour caractéristique d'être très homogènes entre eux et d'être le plus souvent catégoriques, c'est-à-dire qu'ils ne pratiquent (presque) pas la variation. Le profil le plus proche des témoins est celui de personnes âgées parlant la variété arli, ou bien un jeune néo-locuteur arli. Le profil opposé est celui d'individus plutôt jeunes, parlant la variété mečkar, čergar I ou bien arli & čergar I ou bien arli. On constate qu'il y a beaucoup plus d'individus catégoriques de la postposition de l'adjectif que d'individus catégoriques de l'antéposition de l'adjectif. C'est significatif d'une variation de la place de l'adjectif qui tend peut-être à disparaître au profit d'une préférence pour un adjectif fixe, postposé au nom.

On peut remarquer que presque tous les individus catégoriques sont membres des mêmes familles : F65A et M52A sont frères et sœur, et F65A est mariée avec M71A, M52A est marié avec F53A. Ils sont oncles et tantes de F22AČ, F16AČ et M13AČ, qui sont sœurs et frère, ainsi que de M37A, F20A et M11A, cousins entre eux et par rapport aux trois derniers. M18Č est cousin de F22AČ, F16AČ et M13AČ de l'autre côté de la famille. F11A et F9A sont cousines. F47M est mariée avec M45M, et M19M$_{(2)}$ est leur fils. Quant aux autres, F22M, M19M$_{(1)}$ et M22a sont amis, et F37MB, F66M et M58Č sont amis. Un facteur important d'explication de la variation, comme de l'absence de variation, est donc le lien familial, amical ou professionnel, qui amène des individus à se fréquenter de manière récurrente, et donc à partager des innovations et/ou rétentions, voire une même attitude linguistique, au-delà des différences qui existent par ailleurs dans les variables sociales « classiques » telles que GENRE, CLASSE D'AGE, EDUCATION, etc. Nous concevons ce type de lien affinitaire comme la notion anglo-saxonne de pair, *peer*, à ceci près que les *peers* forment un groupe social et cohérent à part entière, partageant les mêmes caractéristiques sociales (appartenant donc aux mêmes niveaux des variables citées ci-dessus). Il serait nécessaire d'en faire une nouvelle variable d'étude. Cependant il nous est impossible de l'intégrer, en raison des conditions de constitution de l'échantillon de participants : celle-ci s'est réalisée par le bouche à oreille, précisément par le biais du réseau des pairs. L'ensemble des participants forme donc une grande chaîne de liens, et rares sont les participants qui ne se connaissent pas entre eux. Cela explique pourquoi les *outliers* qui ont des comportements atypiques ont des liens proches d'un même ordre, et pourquoi leur atypisme est également du même ordre (ils sont *outliers* des mêmes variables).

Enfin, nous avons étudié les individus en fonction de leur âge numérique. L'âge peut en effet s'analyser de deux manières : par tranches (par *classes* d'âge) et par nombre (de manière numérique, avec l'âge des locuteurs en années). On parle, dans le premier cas, de variable catégorielle, et dans le second cas de variable numérique/continue. L'âge numérique ne semble pas être un facteur très pertinent en ce qui concerne l'usage de la polydéfinitude (DEF N DEF A) versus de la monodéfinitude (DEF A N et DEF N A), car ces structures sont réparties de manière équilibrée dans tout l'éventail des âges. On trouve DEF N A plutôt chez les plus jeunes – la frontière se situant entre 19 et 22 ans, et DEF A N plutôt chez les personnes moyennement et plus âgées. D'une manière globale, la postposition de l'adjectif est produite par des personnes de tous âges, mais est hégémonique chez les locuteurs de moins de 22 ans. On peut donc en déduire que, pour la variable CLASSE D'AGE, une frontière d'âge pertinente serait plutôt 20 ans que 15 ans pour distinguer les deux premières modalités.

3.7.5 Synthèse

Les deux versions du test réalisées ont donné des résultats quantitativement, mais non qualitativement différents : ils sont plus contrastés dans le test 1 que dans le test 2. Cependant, nous avons considéré pertinent de faire la moyenne des deux tests. Le test visait à amener les locuteurs à opposer des objets semblables mais différant par une ou deux propriétés qu'ils pouvaient exprimer par des adjectifs. Pour exprimer ces oppositions, il s'agissait de mettre en valeur l'adjectif, fonction qui est assurée par trois stratégies syntaxiques possibles en romani :

- polydéfinitude avec postposition de l'adjectif (Mečkars, Čergars I, 2 classes d'âge moyennes, femmes)
- monodéfinitude avec postposition de l'adjectif (moins de 15 ans, Čergars I, Arlis & čergars I)
- monodéfinitude avec antéposition de l'adjectif (plus de 60 ans, témoins, Arlis, personnes moins scolarisées)

Nous avons également observé la position de l'adjectif dans d'autres syntagmes que ceux introduits par un article défini. Il est remarquable que le profil des locuteurs pratiquant beaucoup DEF N A soit le même que celui de ceux pratiquant beaucoup la postposition de l'adjectif, et pareillement pour DEF A N et l'adjectif antéposé. Le phénomène que l'on trouve dans les syntagmes définis excède donc leur champ et concerne tous les syntagmes nominaux de la langue, quel que soit le déterminant.

4. Synthèse générale

Nous pouvons poursuivre, ici en violet, notre réponse à la problématique (Figure 87) :

Questions structurelles

1) Quels éléments peuvent entrer dans un syntagme polydéfini romani ?

La structure polydéfinie est constituée d'un nom précédé d'un article défini (éventuellement précédé d'une préposition) et d'un adjectif précédé d'un article défini (éventuellement précédés d'une préposition). Un déterminant peut venir remplacer l'article défini introduisant le nom : le possessif (POSS N DEF A). Le romani ne se distingue pas ici du grec.

Le corpus avec rupture interne confirme qu'il est possible d'avoir un possessif à la place du premier article défini (POSS N DEF A), mais également qu'il est possible que deux déterminants introduisent le nom, article défini et quantifieur qui lui succède (DEF QUANT N/ DEF A).

Le corpus semi-spontané montre qu'un autre déterminant peut venir précéder l'article défini en tête de syntagme : le démonstratif (DEM DEF N DEF A). Comme la polydéfinitude se construit exclusivement avec l'article défini, si l'on excepte un cas de répétition du démonstratif, mais jamais avec l'article indéfini, on peut parler d'une variable linguistique de DEFINITUDE.

2) Dans quel ordre les trouve-t-on ? Existe-t-il des structures DEF A DEF N ?

Le corpus spontané présente des syntagmes avec adjectif postposé DEF N DEF A, et quelques cas d'ordre DEF A DEF N. Le corpus avec rupture interne ne présente aucun cas de syntagme polydéfini avec adjectif antéposé.

Dans le corpus semi-spontané, l'ordre est DEF N DEF A avec adjectif postposé dans tous les cas sauf deux : un syntagme DEM DEF A DEF N et un syntagme DEF A/ DEF N avec pause interne. Le romani se distingue donc là du grec. On peut parler d'une variable linguistique d'ORDRE DES CONSTITUANTS préconisant la postposition de l'adjectif articulé.

3) L'ordre des mots est-il plus souple en syntagme poly- que monodéfini ? Le romani présente-t-il la même souplesse que le grec ?

Le romani albanais permet l'antéposition de l'adjectif (DEF A N) aussi bien que sa postposition (DEF N A) en syntagme monodéfini. Il se distingue par là du grec, qui ne permet que l'antéposition de l'adjectif (DEF A N, *DEF N A). Le romani est donc plus souple que le grec en syntagme monodéfini.

En syntagme polydéfini, DEF A DEF N comme DEF N DEF A sont possibles en romani. Toutefois, l'antéposition de l'adjectif semble marginale par rapport à sa postposition, car nous avons relevé

très peu de cas d'antéposition (DEF A DEF N). Le romani se distingue par là du grec, qui antépose volontiers l'adjectif (DEF A DEF N) autant qu'il le postpose (DEF N DEF A). L'ordre des mots est donc moins souple en syntagme polydéfini que monodéfini en romani – c'est le contraire en grec.

4) Lorsqu'une expression polydéfinie contient plusieurs adjectifs, trouve-t-on les mêmes contraintes d'ordre qu'en grec ?

Le corpus spontané ne présente pas de cas de succession d'adjectifs. Tous les syntagmes polydéfinis avec rupture présentent également un seul adjectif.

Le corpus semi-spontané présente quelques cas de succession d'adjectifs. Lorsqu'une expression polydéfinie contient plusieurs adjectifs, ils peuvent encadrer le nom ou bien être tous postposés. S'ils encadrent le nom (DEF A N A et DEF A N DEF A), c'est l'adjectif postposé qui porte le contraste. Il permet de réaliser la sélection dans un ensemble de référents, plutôt que l'adjectif antéposé. Chez les locuteurs anciens et Arli, on a pu constater cependant que l'adjectif antéposé peut également réaliser la sélection d'exemples.

Si les adjectifs sont tous deux postposés, ils sont alors tous deux articulés (DEF N DEF A DEF A), ou bien seulement le premier (DEF N DEF A A), ou seulement le deuxième (DEF N A/ DEF A). Ceci contredit l'affirmation d'Igla (1996 : 166) (cf. section 2.1 du chapitre I, p. 142), selon laquelle « *enthält die NP einen definiten Artikel, muß dieser bei Nachstellung des Adjektivs wiederholt werden* ». C'est le premier adjectif prononcé, soit le plus proche du nom, qui permet d'effectuer la sélection. Nous pouvons donc affirmer que la structure est récursive.

5) L'adjectif d'un syntagme polydéfini peut-il avoir son propre complément ?

Aucun adjectif du corpus spontané n'a son propre complément. Les syntagmes avec rupture n'en présentent pas non plus.

Aucun adjectif du corpus semi-spontané n'a son propre complément.

6) L'emplacement postnominal est-il la zone des « options » ? Quel est le degré de nominalité des éléments qui s'y trouvent ?

L'analyse du corpus semi-spontané montre que la postposition de l'adjectif, même en syntagme polydéfini, sert de position préférentielle pour les adjectif empruntés – qu'ils l'aient été à une langue antéposant ou postposant l'adjectif, qu'ils aient été morphologiquement adaptés ou non, quelle que soit l'époque de l'emprunt. En ce sens, le romani contemporain a compartimenté sa syntaxe nominale (section 2.4.3, p. 335 et 2.4.7, p. 350).

Nous avons constaté que la postposition de l'adjectif est non-marquée dans le contexte de la tâche, où l'adjectif est restrictif. Il s'agit d'une rupture avec les autres contextes énonciatifs, voire avec les autres variétés, pour lesquelles la position non-marquée de l'adjectif est censée être antéposée.

De même, lorsque plusieurs adjectifs encadrent le nom, c'est l'adjectif postposé qui porte le contraste, ce qui confirme que la position non-marquée de l'adjectif à visée contrastive est

postposée. Il convient donc de parler d'emplacements à usages différents, plutôt que d'emplacements marqués ou non-marqués.

Comme l'adjectif restrictif est nécessaire à l'identification, il n'est nullement une « option » : l'emplacement postnominal n'est donc pas une, ou pas qu'une, zone des options.

7) Existe-t-il une restriction quant au type de nom pouvant entrer dans un syntagme polydéfini ?

Nous avons trouvé dans le corpus spontané presque tous les types de noms : un nom propre et des noms communs, un nom indénombrable et des noms dénombrables, un nom d'événement, des noms abstraits, animés, inanimés. Nous n'avons pas trouvé de nom massif ni de nom générique. Ils ont des poids différents et sont aussi bien hérités qu'empruntés.

Les syntagmes avec rupture ne présentent que des noms communs, dénombrables et animés spécifiques. Ils sont de poids léger et d'origine héritée.

Nous n'avons pas trouvé de contraintes quant au type de nom dans le corpus semi-spontané.

8) Existe-t-il des contraintes morphologiques ou étymologiques quant au type de l'adjectif ?

Le corpus spontané ne présente que des adjectifs hérités. Nous n'avons pu établir de corrélation entre leur poids et leur position dans le syntagme. Les adjectifs qualificatifs des syntagmes avec rupture sont tous hérités.

En corpus semi-spontané, nous n'avons pas trouvé de contrainte de morphologie, de poids ni d'origine étymologique pour l'adjectif.

9) Existe-t-il des contraintes syntaxiques ? L'adjectif doit-il être prédicatif ?

Le corpus spontané ne présente que des adjectifs prédicatifs, à l'exception de *vaver*, autre. Les adjectifs qualificatifs des syntagmes avec rupture sont tous prédicatifs.

La nature du corpus semi-spontané fait que tous les adjectifs sont prédicatifs (à l'exception de « autre »).

10) Existe-t-il des contraintes sémantiques ? Doit-il être restrictif ? Certains adjectifs sont-ils exclus, comme les adjectifs relationnels ou ceux lexicalisés et figés avec le nom ?

Les adjectifs articulés sont orientés vers l'objet et non vers le locuteur, restrictifs et intersectifs (ou intersectifs en apparence subsectifs). Ils permettent une interprétation conjointe. Nous avons toutefois observé plusieurs cas où le corpus polydéfini présente un adjectif non-restrictif, notamment lorsqu'il sert à caractériser, à mettre une étiquette (*epithet*). Dans d'autres cas, on peut parler de restriction seulement si l'on considère un ensemble plus large que celui des seuls référents possibles du nom.

Le corpus spontané présente un cas d'adjectif relationnel et deux cas d'adjectifs semi-lexicalisés avec le nom, ce qui montre qu'ils ne sont pas exclus de la polydéfinitude. Toutefois, dans les deux cas, ils impliquaient un contraste implicite ou explicite.

Les syntagmes avec rupture présentent des adjectifs qualificatifs permettant une interprétation intersective conjointe, tous orientés vers l'objet. Plusieurs sont restrictifs, mais d'autres sont ambigus quant à une interprétation restrictive ou non : un cas de l'adjectif de caractérisation et deux cas de restriction dans un ensemble plus large que les seuls référents possibles du nom.

La nature du corpus semi-spontané fait que tous les adjectifs articulés y sont restrictifs. Ils ont tous une fonction « modificatrice » dans le sens de Nølke (2001) et donnent lieu à une interprétation intersective du syntagme nominal. L'analyse nous a également permis de déterminer la fonction des différentes structures syntagmatiques dans le contexte de la tâche. DEF N DEF A est la structure privilégiée à part égale par les adjectifs de toutes origines. Elle est en revanche la structure privilégiée des adjectifs à lecture intersective en apparence subsective, qui se trouve être hérités (taille, *autre*). Pour comprendre la polydéfinitude, on voit qu'il faut donc combiner une explication par l'origine, la position et la sémantique de l'adjectif. C'est la variable du TYPE D'ADJECTIF.

11) Existe-t-il des contraintes quant au type de déterminant ? Peut-il y avoir répétition d'autre chose que de l'article défini, notamment l'article indéfini ?

Seul l'article défini est répliqué dans notre corpus ; aucun déterminant indéfini ne semble pouvoir être répété. Le possessif peut prendre la place de l'une des occurrences de l'article défini – mais il n'est pas pour autant répliqué. C'est l'article défini qui est alors répété : POSS N DEF A. La structure ne se produit donc qu'en contexte défini : un déterminant défini simple ou complexe en première position et un article défini en seconde position.

Les syntagmes à rupture interne nous amènent à nuancer cette dernière affirmation, puisque l'on peut trouver en première position un déterminant défini associé à un déterminant indéfini, le quantifieur : structure DEF QUANT N/ DEF A.

Seul l'article défini est répliqué dans le corpus semi-spontané, à l'exception d'une occurrence de démonstratif (DEM N DEM A).

12) L'article défini du romani est-il explétif ?

Le corpus spontané montre que l'article défini assume toutes les fonctions présentées par Boretzky (2000b : 44-54). Dans les syntagmes polydéfinis à adjectif antéposé, il revêt un usage sémantique (*knowledge-dependent*) ; dans ceux à adjectif postposé, il revêt aussi bien un usage sémantique que pragmatique (*situation-dependent*).

13) Peut-il y avoir répétition du marquage de cas et de la préposition ?

La polydéfinitude donne lieu à la répétition de la préposition lorsqu'elle a fusionné avec l'article défini. Si un syntagme polydéfini est introduit par une préposition monosyllabique cliticisée à DEF$_1$, on la trouve cliticisée à DEF$_2$, donnant lieu à une structure PREP.DEF$_1$ N PREP.DEF$_2$ A.

Le corpus avec rupture interne montre un cas de réplication des marques flexionnelles : l'un des exemples présente une marque de flexion sur le nom et son déterminant, qui est répétée sur l'adjectif et son déterminant postposés. On en déduit que lorsque le syntagme porte une marque de cas, celle-ci est nécessairement répétée, du moins lorsque l'adjectif est postposé.

Le corpus semi-spontané comprend lui aussi plusieurs cas de répétition des marques flexionnelles et une répétition de la préposition.

14) La partie adjectivale est-elle un cas de modification directe ou indirecte ?

Nous avons analysé la partie adjectivale des syntagmes polydéfinis comme un cas de modification indirecte.

La différence entre modification directe et indirecte est d'ailleurs apparue également dans les syntagmes avec rupture et nous avons analysé la partie adjectivale des syntagmes polydéfinis comme de la modification indirecte.

15) Elle-t-elle un cas d'apposition étroite ?

Si l'on conçoit la différence entre apposition étroite et apposition lâche comme une différence structurelle et non sémantique, alors il existe la même distinction en romani. L'apposition étroite, dont ressortent les syntagmes polydéfinis sans rupture, constitue une seule unité prosodique, portant un seul accent. L'apposition lâche, dont ressortent les syntagmes polydéfinis avec rupture, comprend deux sous-unités prosodiques et porte deux accents nucléaires. Le syntagme polydéfini est une apposition étroite, comprenant une partie nominale et une partie adjectivale.

En revanche, il existe une différence importante entre les syntagmes polydéfinis lâches du romani et ceux du grec moderne. En romani ils peuvent être restrictifs (à l'instar des polydéfinis étroits) alors que ceux du grec ne sont pas restrictifs, si l'on en croit la littérature. Il ne s'agit donc pas, en réalité, du même type d'apposition. Plusieurs de nos exemples constituent un autre type encore que celui des appositions lâches de la littérature : restrictifs comme les polydéfinis classiques (apposition étroite), ils comprennent deux sous-unités prosodiques distinctes comme les polydéfinis lâches (apposition lâche). Pour autant, il ne s'agit pas de deux syntagmes juxtaposés DEF N DEF N où l'adjectif est substantivé. En effet, lorsque l'adjectif substantivé est introduit par un article défini, il n'est alors pas restrictif. Lorsque l'adjectif substantivé est restrictif, il n'est alors pas introduit par un article défini, mais par un démonstratif.

16) Existe-t-il des syntagmes polydéfinis lâches ? Leur interprétation est-elle non-restrictive ?

Il existe des syntagmes polydéfinis lâches en romani comme en grec. Ils présentent une rupture entre la sous-partie « nominale » et la sous-partie « adjectivale » : en général une pause, avec parfois un ensemble de mots insérés. Il peut s'agir d'un ou de plusieurs mots : nom propre, interjection, proposition relative, voire le reste de l'énoncé. Cette diversité rappelle celle du grec. L'un de nos exemples montre que le syntagme peut également ne pas présenter de pause.

Dans notre corpus, l'interprétation des syntagmes polydéfinis avec rupture interne est moins systématiquement restrictive que celle des syntagmes polydéfinis sans rupture interne. Il y a donc une différence avec le grec, où un syntagme avec pause ou insertion implique que l'adjectif apporte une information non essentielle et non-restrictive.

17) Y a-t-il un lien entre polydéfinitude et ordre des constituants nominaux ? Lequel ?

Questions sémantiques

1) Existe-t-il une nuance de sens entre syntagme monodéfini et syntagme polydéfini ?

Les syntagmes polydéfinis donnent l'impression d'être plus spécifiques, probablement parce qu'ils sont restrictifs. Les exemples de syntagmes polydéfinis avec rupture fournis par notre corpus sont en revanche tout aussi spécifiques que les syntagmes monodéfinis avec rupture.

2) *Poly*définitude implique-t-elle *sur*définitude ?

La polydéfinitude n'implique pas, à notre sens, de sur-définitude. On pourrait aller jusqu'à dissocier définitude et article défini, en qualifiant celui-ci d'explétif. Il pourrait s'agir d'un marqueur de focalisation. Les syntagmes avec rupture ne se différencient pas des syntagmes sans rupture.

3) Les syntagmes monodéfinis sont-ils restrictifs, tandis que les syntagmes polydéfinis peuvent être restrictifs ou non-restrictifs ?

En syntagme monodéfini comme en syntagme polydéfini, l'interprétation est soit restrictive soit non-restrictive. Le syntagme polydéfini a pour caractéristique d'effectuer une restriction pragmatique parmi les référents possibles du nom. On peut parler d'une fonction d'identification, qui est la fonction majeure de cette structure. Elle est donc notamment employée lorsque les référents possibles du nom sont plus de deux, et que l'identification est alors considérée comme délicate. Dans les cas (beaucoup moins nombreux) où le syntagme a un effet d'emphase, cette fonction d'identification due à la restriction pragmatique ne semble pas être présente.

Les syntagmes avec rupture ne se différencient pas des syntagmes sans rupture : l'interprétation est soit restrictive soit non-restrictive, que le syntagme soit monodéfini ou polydéfini.

4) Si l'adjectif est antéposé au nom, l'interprétation est-elle seulement restrictive ? Si l'adjectif est postposé au nom, l'interprétation peut-elle être restrictive ou non-restrictive ?

En syntagme polydéfini, lorsque l'adjectif est antéposé, l'interprétation est plutôt non-restrictive ; lorsque l'adjectif est postposé, elle est plutôt restrictive. Mais ce ne semble pas être une règle absolue. En syntagme monodéfini, lorsque l'adjectif est antéposé, l'interprétation est restrictive ou non-restrictive ; lorsque l'adjectif est postposé, elle est plutôt restrictive. Là aussi, il ne s'agit que de tendances. C'est donc la position de l'adjectif plutôt que la réplication de l'article défini qui semble jouer un rôle.

Les syntagmes polydéfinis avec rupture présentent tous un adjectif postposé, et une interprétation soit restrictive, soit non-restrictive. Il est donc difficile d'en déduire une tendance. En syntagme monodéfini, lorsque l'adjectif est antéposé, l'interprétation non-restrictive ; lorsque l'adjectif est postposé, elle est restrictive. La taille du corpus est très restreinte, mais elle confirme l'importance de l'ordre des constituants.

Questions fonctionnelles

1) La polydéfinitude est-elle optionnelle en romani comme en grec ? Si oui, qu'est-ce qui provoque le choix du locuteur ?

Le syntagme polydéfini est optionnel en romani comme en grec.

Les variables linguistiques pertinentes sont le contexte défini (DEFINITUDE), la présence d'un contraste implicite ou explicite et la postposition de l'adjectif (ORDRE DES CONSTITUANTS), l'emploi restrictif ou de caractérisation de l'adjectif (TYPE D'ADJECTIF).

2) S'agit-il d'une variation optionnelle de type *syntactic doubling* (Barbiers 2013 : 7) ?

3) Quelle est la fonction des éléments situés dans l'emplacement postnominal (épithète, apposition, *afterthought*…) ?

Le syntagme polydéfini a pour fonction l'identification. C'est pourquoi il comprend, le plus souvent, un adjectif postposé : la postposition est le lieu du focus, où les sèmes spécifiques sont mis en avant. Nous avons vu qu'un syntagme polydéfini peut avoir d'autres fonctions : précision ou auto-correction, récapitulation, *topic-shift*.

Les syntagmes avec rupture ne se différencient pas des syntagmes sans rupture.

Les éléments situés dans l'emplacement postnominal peuvent être des modifieurs épithètes focalisés, selon le résultat de notre étude, ainsi que des appositions lâches. Cela n'empêche pas d'y trouver également des « options » ayant la fonction d'*afterthought* (Matras 2002 : 166).

4) Le syntagme polydéfini implique-t-il un contraste ?

Le focus implique souvent un contraste, mais celui-ci n'est pas nécessairement explicite. En effet, on a pu constater que l'effet de sélection de DEF N DEF A va de pair avec un contraste *in presentia* (référents du même ensemble mentionnés en discours ou présents dans la situation d'énonciation) ; l'effet d'emphase de DEF N DEF A va de pair avec un contraste *in absentia*.

On remarque que le contraste *in presentia*, qui correspond à l'effet de sélection explicite, se trouve plus dans les syntagmes avec rupture que sans rupture.

L'ensemble des syntagmes polydéfinis du corpus spontané implique un contraste, car la tâche à réaliser a été conçue dans ce dessein.

5) Quel est le statut de l'information qu'il porte ?

Le statut de l'information du nom n'est pas pertinent, dans la mesure où les syntagmes polydéfinis ne transmettent pas nécessairement une information nouvelle, mais aussi une information donnée ou accessible. L'adjectif, lui, est toujours nouveau lexicalement : il apporte donc une information nouvelle, une propriété qui permet l'identification du référent. Les syntagmes avec rupture ne se différencient pas des syntagmes sans rupture.

6) Fait-il plutôt partie du topique ou du focus de la phrase ?

Les syntagmes polydéfinis peuvent faire partie soit du topique, soit du focus de la phrase. Les syntagmes avec rupture ne se différencient pas des syntagmes sans rupture.

7) Implique-t-il un focus sur l'adjectif ? Si oui, de quel type ?

Le syntagme polydéfini implique un focus sur la partie adjectivale du syntagme, soit informatif soit contrastif – sans que l'on ait pu déceler de lien avec la position de l'adjectif. Les syntagmes avec rupture ne se différencient pas des syntagmes sans rupture.

8) Le focus est-il situé sur le la totalité du syntagme polydéfini ou sur sa partie adjectivale ? Dans le deuxième cas, la partie nominale constitue-t-elle la présupposition ?

Il apparaît que c'est la topologie du syntagme nominal qui détermine quelle sous-partie constitue la présupposition, et quelle partie constitue le focus. La place postnominale est une position de focus, ou rhématique, tandis que la position anténominale est une position non-focale, ou thématique. La première est le lieu de la modification indirecte, tandis que la seconde est le lieu de la modification directe. Du moins est-ce le cas dans les variétés mečkar et čergar I, la variété arli

présentant (encore) une distinction moins nette entre les deux. On peut ainsi représenter les quatre structurations possibles d'un syntagme nominal avec adjectif épithète sans rupture interne (Tableau 40). Les syntagmes avec rupture ne se différencient pas des syntagmes sans rupture.

Tableau 40 : Quatre structurations possibles du syntagme nominal avec adjectif épithète

	Adjectif antéposé	Adjectif postposé
Sans répétition ART.DEF	DEF A N *o baro Del, to purane breša* adjectif en rôle thématique (présupposé) article défini restrictif ou non	DEF N A *o pete bare, o djalekto romano* adjectif en rôle rhématique (focus) article défini restrictif ou non
Avec répétition ART.DEF	DEF A DEF N *o tikne o maksemi, o barvalo o Skender* adjectif en rôle thématique (présupposé) article défini restrictif	DEF N DEF A *mo phral o tikno, o filme o bare* adjectif en rôle rhématique (focus) article défini restrictif

Questions socio-linguistiques

1) Tous les locuteurs albanais du romani emploient-ils la structure polydéfinie ?

Tous les locuteurs albanais du corpus semi-spontané emploient la structure polydéfinie… à l'exception d'une personne ayant le profil socio-linguistique suivant : homme, non-militant, âgé de plus de 60 ans, parlant la variété arli (dans sa famille d'origine comme dans sa famille actuelle avec son épouse et ses enfants), ayant un niveau d'instruction très faible (école primaire seulement) et bilingue en grec. Ce locuteur possède toutes les caractéristiques pour antéposer l'adjectif et pratiquer peu la polydéfinitude.

Mais étonnamment, ce locuteur a employé la structure polydéfinie dans d'autres occasions que lors de la tâche semi-spontanée. Notre corpus spontané en recèle deux occurrences : c'est lui qui a prononcé les exemples (158) et (170). On peut donc en déduire que la totalité des locuteurs albanais que nous avons rencontrés et enregistrés emploie la polydéfinitude.

2) Observe-t-on une différence de profil socio-linguistique parmi les locuteurs employant cette structure : différence de genre, d'âge ou de niveau d'éducation ?

Le corpus spontané permet quelques premières constatations quant au profil des locuteurs pratiquant la polydéfinitude. Il s'agit aussi bien d'hommes que de femmes, appartenant à toutes les classes d'âge (sauf les plus jeunes ?).

Le corpus avec rupture confirme que la polydéfinitude est pratiquée par tous les locuteurs du romani d'Albanie : homme et femmes de toutes classes d'âge.

L'analyse du corpus semi-spontané a montré que les variables MILITANTISME et EDUCATION sont moins pertinentes que les autres pour expliquer l'usage de la polydéfinitude. En ce qui concerne le GENRE, on constate que les femmes emploient plus la polydéfinitude que les hommes. Les deux variables les plus importantes sont AGE et VARIETE. Le corpus semi-spontané a confirmé les résultats du corpus spontané, à savoir que la polydéfinitude est peu pratiquée par les plus jeunes. Elle est le plus pratiqué par les adultes ayant entre 15 et 59 ans.

3) Quelles variétés présentent le plus cette structure ?

Dans le corpus spontané, la polydéfinitude est employée par des locuteurs de plusieurs variétés : arli, čergar I, mečkar, arli & čergar I.

Dans le corpus spontané avec rupture interne, la polydéfinitude est employée par des locuteurs de plusieurs variétés : arli, mečkar, čergar I & arli, mečkar & čergar I.

Le corpus semi-spontané confirme qu'il existe une variation dialectale quant à l'usage de la polydéfinitude. Les variétés qui l'emploient le plus sont le mečkar et le čergar I. Celles qui postposent le plus l'adjectif, en syntagme monodéfini ou polydéfini, sont le mečkar et le mečkar bi-variétal.

4) Les locuteurs qui emploient cette structure peuvent-ils être influencés par la connaissance d'une ou plusieurs langues étrangères, et si oui, lesquelles ?

Tous les locuteurs de notre étude ont comme langue seconde l'albanais. Ceux qui emploient la polydéfinitude peuvent ne connaître aucune autre langue, ou bien connaître également : le grec, l'italien, le serbe, le macédonien ou le bulgare. D'autres ont des connaissances scolaires, télévisuelles ou militantes de l'anglais, de l'italien et de l'espagnol.

Le grec est une langue à polydéfinitude qui peut fortement influencer les locuteurs de notre corpus. Toutefois, c'est également une langue qui antépose l'adjectif et prohibe la postposition de l'adjectif (non-articulé), ce qui peut avoir l'influence inverse. L'italien et l'espagnol ne présentent pas de polydéfinitude mais sont des langues où la modification indirecte correspond à la postposition de l'adjectif : elles ont pu jouer un rôle.

En anglais au contraire, la modification indirecte correspond à l'antéposition de l'adjectif. De même, le serbe, le macédonien et le bulgare sont des langues slaves antéposant l'adjectif épithète. La connaissance de ces langues n'a pu favoriser l'emploi de la polydéfinitude. Pour en avoir le cœur net, on peut également effectuer une comparaison avec les témoins, locuteurs natifs du macédonien et du serbe. Leur syntaxe nominale est bel et bien fort différente :

- la polydéfinitude est quasiment inexistante (DEFINITUDE).
- la postposition de l'adjectif est quasiment inexistante (ORDRE DES CONSTITUANTS).
- l'antéposition de l'adjectif est non-marquée dans le contexte contrastif de la tâche.
- lorsque plusieurs adjectifs encadrent le nom, c'est l'adjectif antéposé qui porte le contraste, ce qui confirme que la position non-marquée de l'adjectif à visée contrastive est antéposée.

5) Quel est le profil « type » de la personne employant le plus volontiers la structure polydéfinie ?

La personne qui a le plus de probabilité de produire fréquemment DEF N DEF A est une femme, non-militante, âgée entre 15 et 59 ans et parlant la variété mečkar ou čergar I.

6) Quel est le profil « type » de la personne postposant le plus volontiers l'adjectif ?

Dans le corpus spontané, la monodéfinitude avec adjectif postposé est pratiquée par une plus grande diversité de locuteurs, d'âges et de variétés différentes.

Le corpus semi-spontané confirme qu'il existe une variation dialectale quant au placement de l'adjectif épithète en Albanie. L'analyse montre que la postposition de l'adjectif en syntagme monodéfini est le fait des locuteurs les plus jeunes, plutôt locuteurs des variétés mečkar ou čergar I. Plus exactement, la personne qui a le plus de probabilité de produire fréquemment DEF N A est un homme de moins de 15 ans, en cours de scolarisation (ou moyennement scolarisé s'il est plus âgé), non-militant, parlant čergar I ou čergar I bi-variétal (avec mečkar ou arli), ou encore mečkar. La personne qui a le plus de probabilité de postposer l'adjectif (DEF N A ou DEF N DEF A) est un homme ou une femme de moins de 15 ans, non-militant, parlant mečkar ou mečkar bi-variétal (avec čergar I ou bamlija).

Le placement des adjectifs sur l'échelle d'héritage (ou *inheritedness*, « héritéité ») en fonction de leur origine laisse à penser que la structure la plus innovative est DEF N A, et que DEF N DEF A constitue elle aussi une innovation, pour ainsi dire plus ancienne.

7) Quel est le profil « type » de la personne antéposant le plus volontiers l'adjectif ?

Dans le corpus spontané, la monodéfinitude avec adjectif antéposé est pratiquée quasiment exclusivement par les Arlis et les locuteurs de plus de 40 ans.

Dans le corpus avec rupture interne, la monodéfinitude avec adjectif antéposé est le fait d'un profil beaucoup plus restreint de locuteurs que celle avec adjectif postposé : locuteurs arli d'âge moyen ou avancé dans le premier cas, locuteurs de toutes variétés et de tous âges dans le deuxième cas.

L'analyse du corpus semi-spontané confirme que l'antéposition de l'adjectif, mono- ou polydéfini, est le fait des locuteurs de plus de 40 ans et parlant arli. Plus exactement, la personne qui a le plus de probabilité de produire fréquemment DEF A N a plus de 60 ans, est un homme ou une femme militant, qui parle arli, arli bi-variétal ou est un néo-locuteur arli. La personne qui a le plus de probabilité d'antéposer l'adjectif (DEF A N ou DEF A DEF N) est un homme ou une femme de plus de 40 ans, militante, qui parle arli, arli bi-variétal ou est un néo-locuteur arli.

Questions diachroniques

1) À quand remonte cette structure ? Est-il possible de la dater ?
2) D'où provient-elle ? S'agit-il d'une innovation interne ou d'un « emprunt » ?
3) Quel a été le processus d'acquisition de la structure ?
4) Existe-t-elle dans d'autres variétés du romani ? Constate-t-on des différences ?
5) La syntaxe du syntagme nominal est-elle stable en romani ?
6) La syntaxe du syntagme nominal est-elle stable en romani d'Albanie ?

Questions typologiques

1) La polydéfinitude existe-t-elle dans d'autres langues que le romani et le grec ?
2) Quels sont les différents « types » de polydéfinitude dans les langues d'Europe ?
3) Où se situe le romani dans cette typologie ?
4) Comment nommer au mieux la structure qui fait l'objet de la présente étude ?

Figure 74 : Suite de la réponse à la problématique

Chapitre VI : Discussion

Ce chapitre rassemble les conclusions tirées de notre analyse de corpus et les replace dans le contexte des recherches sur la polydéfinitude en général et sur le romani en particulier. Il s'agit de comprendre quel est l'intérêt, pour les locuteurs, d'employer cette structure, plus exactement quel rôle y jouent l'article défini supplémentaire et la place de l'adjectif. Il élabore un scénario socio-linguistique et un scénario diachronique pour expliquer l'usage actuel de la polydéfinitude en romani d'Albanie. Nous nous attachons également à comprendre d'où provient cette structure en romani, ainsi que la postposition de l'adjectif qui l'accompagne. À quelle époque remontent-elles, quelles variétés sont concernées, pourquoi sont-elles apparues dans la langue ? Un tour d'horizon des langues des Balkans et du reste de l'Europe est nécessaire afin de savoir quelles langues disposent d'un article défini, d'une structure polydéfinie et de la postposition de l'adjectif. Cela nous permettra de dresser une typologie de la polydéfinitude et d'y situer le romani.

1. Origine de la polydéfinitude

1.1 Problématique

Est-il possible de déterminer **quand** le romani d'Albanie a acquis la structure polydéfinie ?

I. Un premier scénario serait qu'elle ait été acquise par le vieux romani (*Early Romani*) puis héritée en romani albanais contemporain. Toutes les autres variétés du romani en auraient hérité, certaines l'ayant par la suite perdue.

II. Un deuxième scénario serait qu'elle ait été acquise par le seul romani balkanique durant l'époque médiévale. Les autres variétés d'Europe ne l'auraient donc pas connue.

III. Un troisième scénario serait qu'elle ait été acquise récemment dans la seule Albanie. Les autres variétés du romani ne l'emploieraient donc pas.

Est-il possible de comprendre **comment** le romani d'Albanie a acquis la structure polydéfinie ?

A. Une première hypothèse serait qu'il s'agisse d'une innovation interne, qui a une fonction particulière et remplit un besoin linguistique.

B. Une deuxième hypothèse serait qu'il s'agisse d'un emprunt à une autre langue. Reste à savoir laquelle ou lesquelles, comment et quand s'est effectué l'empunt.

Pour répondre à ces questions, il est nécessaire de combiner l'histoire de la langue et sa géographie. Nous appliquerons les critères méthodologiques proposés par Adamou (2012 : 61) pour déterminer si le phénomène est dû au contact de langues :

1. Décrire le fait linguistique dans les vernaculaires. (chapitres IV et V)
2. Rechercher le fait linguistique en diachronie. (section 1.2, p. 408)
3. Repérer le fait linguistique dans les autres langues de la même famille. (sections 1.3 p. 411, 1.4.1.1 p. 424 et 1.4.2.1 p. 436)
4. Evaluer la rareté typologique du fait linguistique. (sections 1.4.1.2 p. 425 et 1.4.2.2 p. 438)
5. Identifier la (les) langue(s) source(s) et attester le fait linguistique. (section 1.5, p. 444)
6. Tracer les contacts effectifs entre les locuteurs, région par région. (chapitres IV et V)
7. Identifier le processus. (sections 1.5 p. 444, 2.5 p. 457 et 3 p. 462)

1.2 Histoire : le romani albanais ancien

Il existe peu de textes historiques en romani d'Albanie. Stuart Mann a transcrit plusieurs contes et histoires (de longueurs variables) entre 1933 et 1937, présentés dans le Tableau 94. Le Tableau 95 présente la structure des syntagmes nominaux avec adjectif que nous avons relevés dans ces textes.

On compte quatorze syntagmes définis avec adjectif antéposé, dont 8 monodéfinis (DEF A N), un qui un comprend deux adjectifs successifs (DEF A A N), deux dont le déterminant est un déterminant défini complexe, démonstratif ou possessif (DEM A N ou POSS A N) et trois qui comprennent en réalité deux déterminants : un démonstratif et un article défini (DEM DEF A N). On relève huit syntagmes définis avec adjectif postposé, dont un monodéfini (DEF N A), cinq polydéfinis étroits (DEF N DEF A), un polydéfini large (DEF N/ DEF A) et un qui comprend en réalité deux déterminants : DEF N POSS A.

Mann (1933a : 24) écrit, à propos de l'expression *e jéta e parní*, le monde de la lumière, litt. « la vie blanche », que l'ordre des mots est influencé par l'albanais *jeta e bardhë*, la vie blanche. Ce syntagme comprend en effet deux articles : l'article défini clitique postposé au nom, *jeta*, et l'article de connexion de l'adjectif au nom, antéposé à l'adjectif, *e bardhë*. Il ne fait en revanche aucun commentaire sur les autres syntagmes polydéfinis. Notons que l'on relève dans le texte Ⓓ trois cas d'ordre des mots DEF N A où l'adjectif n'est toutefois pas épithète du nom, mais attribut du sujet, avec élision de la copule. Ce sont des cas de prédication nominale.

(260) Romani (Mann 1934a : 43)

Lék'i	*rovlí*	*arék'i,*	*léki*	*staǵí*	*rupuní,*	*lék'i*	*kundrí*	*balikané:*	*telál*	*serčalí.*
son	bâton	en or	son	chapeau	en argent	sa	chaussure	de porc	en bas	en verre

« Son bâton est en or, son chapeau en argent et ses chaussures en cuir de porc : le dessous est en verre. »

Tableau 94 : Contes collecté par Stuart Mann en 1931

	Titre	Source	Date et lieu	Nom du locuteur	Age	Variété
Ⓐ	*Jek paramič*, un conte	(Mann 1933a : 1-25)	20 juin 1931 à Tiranë	Ćerím, originaire de Shkodër	environ 40 ans	čergar I ? [1]
Ⓑ	*O trin phral*, les trois frères	(Mann 1933a : 26-32)	27 avril 1931 à Tiranë	Delía, fils de Ćerím	15 ans	čergar I ? [1]
Ⓒ	*O zylkanôni thai e lačí e devlkék'i*, le favori et la sainte	(Mann 1933b : 147-152)	3 juin 1931 à Tiranë	Imíni, originaire de Shkodër	22 ans	čergar I ? [1]
Ⓓ	*O ǵyćél thai e máčka*, le chien et le chat	(Mann 1934a : 34-45)	12 avril 1931 à Tiranë	Ismáil Abazi, originaire de Shkodër	10 ans	čergar I ? [1]
Ⓔ	*So ćerél o χaló*, les aventures du chauve	(Mann 1934b : 153-158)	22 juin 1931 à Tiranë	Ćerím	environ 40 ans	čergar I ? [1]
Ⓕ	*E ǧuváći morći*, la peau de puce	(Mann 1941 : 85-87)	5 avril 1931 à Tiranë	Ismáil Abazi	10 ans	čergar I ? [1]
Ⓖ	*O ralkó thai o arápo thai i raklí*, le garçon non-rom, le noiraud et la fille non-rom	(Mann 1935 : 177-184)	mars 1931 à Tiranë	Neǧípi, originaire de Thessalonique	environ 40 ans	čergar II ? [2]
Ⓗ	*O sap so k'erdilo manúš*, le serpent qui se fit homme	(Mann 1937 : 186-192)	1931 à Tiranë	Neǧípi et Sali	environ 40 ans, 32 ans	čergar II ? [2]
Ⓘ	*Jek masáli*, une histoire	(Mann 1937 : 192-195)	1931 à Tiranë	Neǧípi et Sali	environ 40 ans, 32 ans	čergar II ? [2]

[1] L'auteur qualifie cette variété « variété de Scutari ou du nord de l'Albanie ».

[2] L'auteur qualifie cette variété « variété de Korça ou du sud de l'Albanie ou de Macédoine occidentale ».

Tableau 95 : Relevé des SN définis avec adjectif épithète dans les contes collectés par Stuart Mann

Texte	(PREP) DEF A N	DEM DEF A N	DEM A N ou POSS A N	DEF A A N	(PREP) DEF N A	(PREP.)DEF N POSS A	DEF N DEF A
Ⓐ	*o tsiknó manúš*, le petit homme	*okova o tsiknó manúš*, ce petit homme		*o tsiknó divó manúš*, le petit homme fou	*an(d)é jéta parní*, dans le de la lumière (3)	*po šoró léhko čhindó*, sur sa tête coupée	*o rakló o χaló*, le garçon non-rom chauve
	o maškarúno phral, le frère du milieu (2)	*gov o divó manúš*, cet homme fou					*andé jéta e parní*, dans le monde de la lumière
	e maškarúne phéja, les sœurs du milieu						*o ôrló – o parnó –* l'aigle/ blanc
	andé kalí jéta, dans le monde des ténèbres						
	e parné dašošé (šəŋg), (la corne) du bélier blanc						
Ⓑ	*o bilačó manúš*, le méchant homme	*okôva o bilačó manúš*, ce méchant homme	*kálé bilačé manušé*, à ce méchant homme				*me phralék'i e barék'i*, de mon grand frère
							ano pai o ratvaló, dans l'eau ensanglantée
Ⓓ	*o Déti Kálo*, la Mer Noire		*lék'e bal baré*, ses poils longs				*o Déti o Kálo*, la Mer Noire
Ⓖ	*o tiknó arápo*, le petit noiraud						

Ⓒ, Ⓔ, Ⓕ, Ⓗ Ⓘ : Nous n'avons pas relevé de syntagme défini avec adjectif épithète dans ces textes.

Ainsi constate-t-on que la polydéfinitude existait déjà en romani albanais des années 1930. La postposition de l'adjectif en syntagme monodéfini existait également, mais semble moins usitée. L'adjectif était fréquemment antéposé, en syntagme monodéfini. Les locuteurs interrogés par Stuart Mann vivaient à Tiranë mais étaient originaires de régions différentes (Shkodër et Korçë), leur famille provenant respectivement de « Serbie de sud » (l'actuel Kosovo), et du nord de la Grèce et de Macédoine. La zone d'extension de la polydéfinitude correspondait au minimum à la zone triangulaire représentée en rouge sur la Figure 75 : Albanie, Kosovo, la majeure partie de la Macédoine historique (slave, grecque et albanaise).

Figure 75 : Zone d'extension de la polydéfinitude en romani d'Albanie en 1931

1.3 Géographie : le romani des autres régions

1.3.1. Descriptions d'autres variétés

Dans quelles variétés de la langue trouve-t-on la structure polydéfinie ? Nous avons vu en section 2.1 du chapitre I, p. 142, qu'elle a été relevée par les travaux suivants :

- Gjerdman et Ljungberg (1963) : kalderaš de **Suède**
- Boretzky (1993 : 171) : bugurdži du **Kosovo**
- Boretzky (1994) : kalderaš du Banat (où se parle notamment le valaque) en **Serbie**
- Igla (1996 : 40, 85) : variété *vlax* d'Ajia Varvara en **Grèce**
- Deman (2005 : 73) : variété *vlax* de Bogota en **Colombie**
- Matras (2004 : 74-75): variété de Parakalamos en Grèce

La présence de la polydéfinitude dans de si nombreuses variétés invalide notre **hypothèse III** selon laquelle la polydéfinitude ne serait apparue que dans la seule Albanie. Les sections suivantes présentent des données et des publications comprenant des syntagmes polydéfinis.

1.3.2. Le corpus RMS

Un moyen de comparer les variétés est de consulter le corpus en ligne *Romani Morpho-Syntax Database*. Nous avons trouvé des syntagmes polydéfinis[1] dans des variétés d'Albanie, de Grèce, de Serbie, de Moldavie et de Roumanie (Tableau 96).

[1] Nous avons exclu du repérage les cas de syntagmes nominaux avec adjectif au superlatif.

Tableau 96 : Variétés du corpus RMS comportant de la polydéfinitude

Pays	Ville	Variété	Code	Occurrences
Albanie	Fier	Mečkari	AL-001	2
Grèce	Parakalamos	Romalicikanes	GR-002	1
Yougoslavie	Beočin	Arli	YU-011	2
Moldavie	Chișinău	Čurari	MD-006	1
	Zirnești	Laeši Kurteja	MD-007	2
Roumanie	Maglavit	Ursari	RO-004	1
	Pitești	Spoitori	RO-006	2
	Pitești	Kaldaraš	RO-008	1
	Huedin (enregistré à Grenade)	Piculesči	RO-013	2
		Kurturare	RO-015	3
	Călărași (enregistré à Grenade)	Spoitari	RO-016	3
	Deaj	Kərəmidarea	RO-025	1
	Senereuș	Maj Vlaši	RO-058	2
	Țăndărei	Kangljari/Peptenari	RO-064	1
	Timişoara	Kalderaš	RO-065	1
	Țăndărei	Kangljari/Peptenari	RO-066	1

Le Tableau 97 liste les phrases du questionnaire concernées ainsi que leur traduction dans les différentes variétés. Même s'il s'agit d'un corpus non-spontané, il fournit un premier indice quant à la dispersion géographique de la polydéfinitude. On constate que l'Albanie est loin d'être le seul pays où s'emploie la polydéfinitude, et que la Roumanie est le pays qui l'emploie le plus largement.

Les variétés de Roumanie et de Moldavie présentent une grande diversité de structures :

- à deux déterminants avec réplication de l'article défini (DEF N DEF A)

  + en romani de Serbie (Voïvodine), réplication du possessif : POSS N POSS A
- à deux déterminants : possessif associé à l'article défini (POSS N DEF A), quantifieur associé à l'article défini (QUANT N DEF A)
- à trois déterminants : réplication de l'article défini associée à un possessif (DEF N POSS DEF A)
- à quatre déterminants : POSS N DEF NUM A et sa variante POSS NUM N DEF A
- à quatre déterminants avec réplication du possessif : POSS NUM POSS N DEF A

  + en romani de Grèce, à quatre déterminants avec réplication de l'article défini : DEF NUM DEF A POSS N

Tableau 97 : Syntagmes polydéfinis dans le corpus RMS

Variété	Phrase	Phrase en anglais	Traduction en romani	Structure
Mečkari	480	*My little sisters, I saw you when you were eating sweets!*	*Mo phenja o tikine dikhlom tumen ka xalen o karamele*	POSS N DEF A
	513	*I saw my (two) younger sisters exactly a year ago.*	*Dikhlom mi phenjen e tikinen akate ta jek bərš*	POSS N DEF A
Čurari	691	*There is a big celebration in our village every New Year's Eve.*	*Ande firiso brš o nevo si bari srbutorja ande gav*	QUANT N DEF A
Laeši Kurteja	513	*I saw my two younger sisters exactly a year ago.*	*Me dikljem mre dujen me phejange e cinorenge ek brš palpale*	POSS NUM POSS N DEF A
	-	*Now I'm studying psychology and I'm also a media reporter a journalist for the Romani programme Petalo Romano.*	*Akanak sikavav pe psihalogija haj maj sim informatoro masmedija žurnalisto pa i programa i romani Petalo Romano.*	DEF N DEF A
Ursari	652	*My little daughter is scared to go across a bridge.*	*Me čhajake le tiknjake si lake traš te nakhel o podos*	POSS N DEF A
Spoitori	513	*I saw my (two) younger sisters exactly a year ago.*	*Dikhom me pen o cinen ekzakto akana ek breš*	POSS N DEF A
	652	*My little daughter is scared to go across a bridge.*	*E čhaj miri e cikni trašala lako zi te nakel o podos*	DEF N POSS DEF A
Kaldaraš	652	*My little daughter is scared to go across a bridge.*	*Myrə šejakə la ternjakə sy lakə dar te nakhəl o podo/ i phurt*	POSS N DEF A
Piculesči	380	*He saw the old man.*	*Vo dikhle le romes le phures.*	DEF N DEF A
	691	*There is a big celebration in our village every New Year's Eve.*	*Sy ek bari festa ne amaro gav se no borš o njevo.*	DEF N DEF A
Kurturare	380	*He saw the old man.*	*Vo dikhle manušes le phures.*	(DEF) N DEF A
	480	*My little sisters, I saw you when you were eating sweets!*	*Məri phej e cini, dikhlem tumen kana xanas guglimata.*	POSS N DEF A
	652	*My little daughter is scared to go across a bridge.*	*Məri šej e cini sy lakə dar te nakel o podo.*	POSS N DEF A
Spoitari	480	*My little sisters, I saw you when you were eating sweets!*	*Me phe o cine dikhom tumen kibor gudimata xane.*	POSS N DEF A

Variété	Phrase	Phrase en anglais	Traduction en romani	Structure
	513	*I saw my two younger sisters exactly a year ago.*	*Me dikhom me phen o do cignen ekzaktos o breš ka nakas.*	POSS N DEF NUM A
	652	*My little daughter is scared to go across a bridge.*	*Me rake kəzi o cine ko trašal te nakel o paj o podos*	POSS N [?] DEF A
Kərəmidarea	513	*I saw my two younger sisters exactly a year ago.*	*Dikhljom myrə do phenjan le cinonjan akana ek bəš.*	POSS NUM N DEF A
Maj Vlaši	652	*My little daughter is scared to go across a bridge.*	*Mure šaorake la cyne ke dar hi te nakhəl o podo.*	POSS N DEF A
	691	*There is a big celebration in our village every New Year's Eve.*	*Syn ek baro serbatoare ade savro data ko bərš o nevo.*	DEF N DEF A
Kangljari/Peptenari	652	*My little daughter is scared to go across a bridge.*	*I čhej mury i cəni si lake dar te nakhel o podo.*	DEF N POSS DEF A
Kalderaš	652	*My little daughter is scared to go across a bridge.*	*Murra šej la ciknja lakə dar te nakel o podo.*	POSS N DEF A
Kangljari/Peptenari	652	*My little daughter is scared to go across a bridge.*	*Lake dar te nakhel o podo murre čhej i cikni.*	POSS N DEF A
Romalicikanes	513	*I saw my two younger sisters exactly a year ago.*	*Dikhljomas e dujen e čikoren mi phenjen prin nje berš.*	DEF NUM DEF A POSS N
Arli	691	*There is a big celebration in our village every New Year's Eve.*	*Ana maro gav isi bari proslava svako berš ko dive e beršeso o nevo.*	DEF N DEF A
	480	*My little sisters, I saw you when you were eating sweets!*	*Me phenja me tikne dikhljem tumen kaj haljen gudljipe.*	POSS N POSS A

1.3.3. Données roumaines contemporaines

Nous avons étudié les quatre premiers récits de vie de *Romane Asva* (Tirard 2014), publiés sous forme de livre (Cioabă 2006a) et de DVD (Cioabă 2006b). Il s'agit de témoignages de femmes et d'hommes ayant subi la déportation durant la Seconde Guerre Mondiale. Tous parlent des variétés *vlax* :

- Premier récit : Mello, homme vivant à Sebeş, déporté en Olténie (enregistré le 2 juin 2004),
- Deuxième récit : Lenuţa, femme vivant à Galaţi,
- Troisième récit : Traian Grancea, homme de 79 ans né et vivant à Porumbacu de Jos, déporté en Transnistrie (enregistré le 22 juin 2004),
- Quatrième récit : Costea, homme vivant à Ştefanciu Mare,
- Intervieweuse : Luminiţa Mihai Cioabă, femme de 47 ans, née à Târgu Cărbunești et vivant à Sibiu.

Nous avons observé la place de l'adjectif et des déterminants dans l'ensemble des syntagmes nominaux avec adjectif épithète. Les occurrences de polydéfinitude sont les suivantes :

- trente-quatre cas de syntagme polydéfini avec adjectif postposé (**DEF N DEF A**), avec éventuelle réplication du morphème flexionnel ou de la préposition : *le love le rumïnìsko*, l'argent roumain, *le thema ol bravale*, les pays riches, *le pòli le sumnakune*, les pièces d'or, *le ćhejan le barean*, les grandes filles, *anθa-l rom le aśunde*, de Roms célèbres, *o Parolea o phuro*, Parolea le vieux, *anθ-o paj o śudro*, dans l'eau froide, *p-o fròntо o rumïnìsko*, sur le front roumain, *le gàlbi ël barë*, les grosses pièces d'or, *le ćhaworrënça le cïnorënça*, avec les petits enfants, *le Iliasqo le phurësqo*, d'Ilias le vieux, *ka-o spitàlo o romano*, à l'hôpital rom,
- un cas de syntagme polydéfini avec adjectif antéposé (**DEF A DEF N**) : *o rumïno o rëzbòj*, la guerre roumaine,
- deux cas de syntagme polydéfini introduit par un possessif (**POSS N DEF A**) : *tumarë pàpuri le phurë,* vos vieux grands-parents,
- cinq cas de syntagme polydéfini avec un déterminant défini complexe supplémentaire (**DEF N POSS DEF A, DEF N DEM DEF A, DEF N DEF A DEM**) : *e tradìcia amari e romaii*, notre tradition rom, *le romnea kodola la phurea*, ces vieilles femmes, *ol thema ol strëine kodola*, ces pays étrangers, *anθ-o khër o pustìo kodowa*, dans cette maison déserte.

On observe des points communs entre ces récits de vie et de notre corpus albanais :

- l'adjectif est quasiment toujours postposé,
- si le premier article est fléchi, le deuxième l'est aussi : *le love le rumïnsko, le ćhejan le barean, ol thema ol strëine kodola, le ćhaworrënça le cïnorënça, le Iliasqo le phurësqo,*
- la marque de flexion nominale est répliquée sur l'adjectif : *le ćhejan le barean, le romnea kodola la phurea, le ćhaworrënça le cïnorënça, le Iliasqo le phurësqo,*
- le premier déterminant peut ne pas être un article défini, mais un autre déterminant défini : *tumarë pàpuri le phurë,*

- la préposition qui fusionne avec le premier article défini n'est pas nécessairement répliquée sur le deuxième : *p-o frònto o rumïniìsko*,
- si le premier article défini est introduit par une préposition monosyllabique non fusionnée avec l'article ou une préposition plurisyllabique, la préposition n'est pas répliquée : *ka-o spitàlo o romano, anθ-o khër o pustìo kodowa, anθ-o paj o śudro, anθa-l rom le aśunde*.

On peut se demander si la réplication de la préposition fusionnée est possible dans cette variété. Nous pensons que c'est le cas, car on la trouve dans des syntagmes sans adjectif, où la deuxième instance de la préposition (même longue) « accompagne » un déterminant démonstratif : *anθ-ël rïpe anθ-ël kodola*, dans ces ravins, littéralement « dans les ravins dans les ces » (structure très complexe PREP DEF N PREP DEF DEM). La préposition est également répliquée dans les appositions étroites nominales : *ka-o pòdo ka-o Nistro*, au pont du Dniestr, littéralement « au pont au Dniestr ».

On constate une divergence par rapport à notre corpus albanais : l'article défini répliqué peut subir quelques « déformations » phonologiques, notamment vocaliques : *le thema ol bravale, anθa-l rom le aśunde, le romnea kodola la phurea, le gàlbi ël barë*. On constate un autre phénomène : un déterminant défini complexe peut venir s'ajouter au syntagme, qui a alors trois déterminants (*e tradìcia amari e romaii, le romnea kodola la phurea, ol thema ol strëine kodola, anθ-o khër o pustìo kodowa*). Cette variété de romani pratique volontiers l'accumulation de déterminants, même en l'absence d'adjectif :

- DEF DEM DEF N : *o kodowa o ćhaw'*, ce garçon (litt. « le ce le garçon »),
- DEF DEM POSS N : *le kodole amarë romes*, nos Roms/maris (litt. « les ces nos Roms »),
- DEF N DEF DEM : *le kurùza le kodola*, ces céréales (litt. « les céréales les ces »), *ël magàrea le kodola*, ces chevaux (litt. « les chevaux les ces »), *anθ-ël rïpe anθ-ël kodola*, dans ces ravins (littéralement « dans les ravins dans les ces »),
- DEF DEM POSS N : *sea-l boàle kadala*, toutes ces maladies (litt. « toutes les maladies ces »).

1.3.4. Données roumaines anciennes

Un recueil de textes en romani de Valachie et de Moldavie, recueillis au milieu du XIXème siècle, a été publié en 1878 par le folkloriste Barbu Constantinescu. Il s'agit du seul ensemble de textes en langue romani des principautés roumaines du XIXème siècle. Il n'a toutefois pas publié son esquisse de dictionnaire romani-roumain, ni la totalité des contes et des poésies qu'il a collectés dans quatre-vingt-neuf lieux en Valachie et huit lieux en Moldavie. Julieta Rotaru a entrepris d'éditer et de publier le reste de son travail (Constantinescu 2016) (Rotaru 2018). Comme noté par Rotaru (2016), ces textes issus de plusieurs variétés regorgent de réplications de l'article défini. Le Tableau 98 présente quelques occurrences, la traduction donnée par l'autrice en roumain et en anglais, ainsi que notre traduction en français, mais à partir du romani. Cela montre que la structure polydéfinie est répandue de très longue date dans toute la zone roumanophone.

Tableau 98 : Exemples de syntagmes polydéfinis dans les textes collectés par Barbu Constantinescu

Dialecte	Comté	Texte	Structure	romani	roumain	anglais	français
ursari *ricinearísko ursăresc*	Bacău	155	PREP N DEM A	*opr-o plai ko zeleno*	*pe acel deal verde*	*on that green hill*	sur cette colline verte
			DEF N DEF A	*e grasni e jungali*	*iapa cea rea*	*the evil mare*	la méchante jument
	Ploiești	36	PREP DEF N DEF A	*anda-u them o nemțisko*	*din țara nemțească*	*from the German country*	du pays allemand
	Prahova	48	PREP DEF N DEM A	*pașa-i romni ke terni*	*lângă acea nevastă cea tânără*	*next to that young wife*	près de cette jeune épouse
			PREP DEF N DEM A	*andar e kurva ke marime*	*din curva cea spurcată.*	*because of the soiled whore*	de cette catin souillée
		196	PREP DEF N DEM A	*andar o drom ko baro*	*în drumul cel mare*	*on the main road*	sur cette grand-route
laiesi *lăiășisko lăieșesc*	?	43	DEF N DEF A	*i kurva-i marime*	*curva cea spurcată*	*the soiled whore*	la catin souillée
	Ilfov	8	DEF N DEF A	*la juvlea la șukarea*	*pe femeia cea frumoasă*	*the beautiful woman*	la belle femme
			PREP DEF N DEF A	*pala-i juvli e marime*	*după femeia cea spurcată*	*to the soiled woman*	après la femme souillée
		245	PREP DEF N DEM A	*and-o posto ko baró*	*în postul cel mare*	*during the Fast of the Easter*	pendant ce grand jeûne
		84	PREP DEF N DEF A	*pă-l ćereate-l șukareate*	*pe iarba cea frumoasă*	*on the beautiful grass*	sur la belle herbe
		76	PREP DEF N DEF A	*ka-u pelalo o baro*	*la zapciul cel mare*	*to the great tax collector*	le grand *zapciu*
			PREP DEF N DEM A	*and-o posto ko baro*	*în postul cel mare*	*during the lent*	pendant ce grand jeûne
	Damboviţa	131	PREP DEM N DEF A	*pe kou plai o zeleno*	*pe acel deal verde*	*on that green hill*	sur cette verte colline
	Olt	154	PREP DEF N PREP DEF A	*ko şavo ko şukaro*	*acel băiat frumos*	*that handsome boy*	ce beau garçon

Dialecte	Comté	Texte	Structure	romani	roumain	anglais	français
	Teleorman	165	DEM N DEM A	*kol romnea kola p<h>urea*	*acea femeie bătrână*	*that old woman*	cette vieille femme
	Gorj	81	PREP DEF N DEF A	*telau văș o zeleno*	*pe sub pădurea cea verde*	*under the green forest*	sous le bois vert
			DEF N DEF A	*o abeau o romano*	*nunta țigănească*	*Gypsy wedding*	le mariage rom
kalderaš *kikavearisko căldărăresc*	Craiova	21	DEF N DEF A	*o bîrș o baró*	*anul cel mare*	*one great year*	la grande année
	Iași	66	POSS N DEM A	*amaro p<h>ral kou baro*	*fratele nostru cel mare*	*our eldest brother*	notre frère aîné
tismani *tismănarisko tismănăresc*	Gorj	183	DEF N DEM A	*hol șavesa ko kalo*	*cu băiețelul cel negru*	*with the dark chap*	avec ce garcon noir
	Mehedinți	191	POSS N DEM A	*muro pînro ko parno*	*al meu picior cel alb*	*my fair leg*	ma jambe blanche
zlatari *zlătarisko zlătăresc*	București	239	POSS N DEF A	*léski raklí e barí*	*a lui fată mare*	*his big girl*	sa grande fille
			POSS N DEF A	*lésko rakló o ţîknó*	*al lui băiat mic*	*his little boy*	son petit garçon
allemand *nemțisko nemțesc*	?	3	PREP DEF N DEF A	*p-ăi har ăi zeleno*	*pe valea cea verde*	*on the green valley*	dans la verte vallée

On observe des points communs entre la polydéfinitude de ces textes et de notre corpus albanais :

- l'adjectif est toujours postposé,
- si le premier article est fléchi, le deuxième l'est aussi : *la juvlea la șukarea, pă-l ćereate-l șukareate*,
- la marque de flexion nominale est répliquée sur l'adjectif : *la juvlea la șukarea, kol romnea kola p<h>urea, pă-l ćereate-l șukareate*,
- le premier déterminant peut ne pas être un article défini, mais un autre déterminant défini, comme le démonstratif : *pe kou plai o zeleno*, ou le possessif : *léski raklí e barí, lésko rakló o ţîknó*,
- la préposition qui fusionne avec le premier article défini est répliquée sur le deuxième : *ko şavo ko şukaro*,
- la préposition qui fusionne avec le premier article défini n'est pas nécessairement répliquée sur le deuxième : *pă-l ćereate-l șukareate, pe kou plai o zeleno*,
- si le premier article défini est introduit par une préposition monosyllabique non fusionnée avec l'article ou une préposition plurisyllabique, la préposition n'est pas répliquée : *anda-u them o nemțisko, pala-i juvli e marime, ka-u pelalo o baro, telau văș o zeleno*.

On constate également des divergences par rapport au romani d'Albanie :

- l'adjectif peut ne pas être fléchi : *hol șavesa ko kalo*,
- l'article défini répliqué peut subir quelques « déformations » phonologiques, notamment vocaliques : *anda-u them o nemțisko, ka-u pelalo o baro, telau văș o zeleno*,
- le deuxième déterminant peut ne pas être un article défini, mais un autre déterminant défini, comme le démonstratif : *opr-o plai ko zeleno, pașa-i romni ke terni*[1], *andar e kurva ke marime*[1], *andar o drom ko baro, and-o posto ko baró, hol șavesa ko kalo*.

On a vu que le premier comme le deuxième déterminant peuvent ne pas être un article défini, mais un démonstratif ou un possessif. Il arrive même qu'aucun des deux ne soit un article défini : on a alors une structure polydéfinie sans article défini, comprenant seulement des déterminants définis complexes. Notre relevé comporte les structures suivantes :

- POSS N DEM A : *amaro p<h>ral kou baro*, littéralement « notre frère celui grand », *muro pînro ko parno*, littéralement « mon pied celui blanc »,
- DEM N DEM A : *kol romnea kola p<h>urea*, littéralement « cette femme cette vieille ».

Cette dernière structure consiste à réduplicquer le démonstratif au lieu de l'article défini. Rappelons que notre corpus semi-spontané albanais présente lui aussi une structure « polydémonstrative » (exemple (205).

[1] Une autre interprétation de ce syntagme traiterait *ke* non comme un démonstratif, mais comme la préposition *k/ke*, à. Elle « remplacerait », dans la sous-partie adjectivale, toute préposition introduisant le syntagme nominal – ce qui nous semble très étrange. Une telle interprétation pourrait être étendue aux syntagmes masculins (*opr-o plai ko zeleno*).

Ces données prouvent que le romani de Roumanie connaît la structure polydéfinie depuis plus d'un siècle et demi et qu'il a innové très tôt dans l'usage des prépositions et de la flexion. Il présentait des structures polydéfinies très complexes et pouvait même répliquer le démonstratif. Gardons toutefois en mémoire que ces occurrences sont issues de chants, donc d'une parole non-spontanée, élaborée, soumise aux impératifs de la métrique.

1.3.5. Autres données

1.3.5.1 Macédoine et Kosovo

Souvenons-nous que l'un des témoins de Macédoine que nous avons pu enregistrer pour notre corpus semi-spontané a produit un syntagme polydéfini avec rupture interne (et adjectif postposé), exemple (257) – mais pas l'autre témoin de Macédoine. Le témoin du Kosovo a quant à lui produit deux syntagmes polydéfinis, l'un avec rupture et l'autre sans rupture (et adjectif postposé) (cf. section 2.3.2.5 du chapitre V, p. 324, exemples (258) et (259)). Ainsi la polydéfinitude est-elle possible en Macédoine et au Kosovo, mais très rare, et ce même dans le contexte contrastif de la tâche.

1.3.5.2 Grèce

Larissa

J'ai pu interviewer plusieurs Roms au cours d'une mission de terrain à Larissa (Thessalie) à l'été 2014. Mon hôte, un homme de quarante-et-un ans parlant une variété *vlax* (présentant également des traits balkaniques I) a prononcé plusieurs syntagmes polydéfinis, avec adjectif postposé et antéposé :

(261) *M41VlGr, Larissa, 27 août 2014*

Asun![1]	*Amen*	*kate*	*daa*	*but*	*borbisar-as/*	
écouter.IMP	1.PL	ici	encore	beaucoup	parler-1PL	
i	*t͡shib*	*i*	*romanik-a.*	*And-i*	*Larissa*	*kate.*
ART.DEF.F.SG	langue(F.SG)	ART.DEF.F.SG	rom-F.SG	à-ART.DEF.F.SG	Larissa	ici

« Ecoute ! Nous ici, on parle encore beaucoup la langue romani. Ici à Larissa. »

(262) *M41VlGr, Larissa, 27 août 2014*

Nastin	*akhar-esa*	*e*	*vorb-es*	*e*	*roman-es/*	*ka*	*phen-av*
ne_pas_pouvoir	comprendre-2SG	ART.DEF.PL	mot-F.PL	ART.DEF.PL	rom-F.PL	FUT	dire-1SG

[1] La lettre *s* représente le son [s] de l'Alphabet Phonétique International : le locuteur ne prononce pas de fricative postalvéolaire [ʃ], mais une fricative alvéolaire rétractée. Il ne prononce pas non plus d'affriquée postalvéolaire (éventuellement aspirée) [t͡ʃʰ], mais une affriquée alvéolaire rétractée (éventuellement aspirée) [t͡sʰ].

tu-ke.	*Nas̱tin*	*akhar-esa*	*len*	*sa.*
2SG.OBL-DAT	ne_pas_pouvoir	comprendre-2SG	3.PL.OBL	tout

« On ne peut pas comprendre tous les mots romani/ je vais te dire. On ne peut pas tous les comprendre. »
(parlant de ses difficultés à comprendre la variété parlée par son épouse)

(263) *M41VlGr, Larissa, 27 août 2014*

O	*Del*	*prep/*	*te*	*phabar-el*	*len*	*olen.*	*Te*
ART.DEF.M.SG	Dieu(M.SG)	il_faut	COMPL	brûler-3SG	3.PL.OBL	3.PL.OBL	COMPL
astar-el	*e*	*s̱ukar*	*manus̱-es/*	*e*	*sukar*	*e*	*Rom-a.*
attraper-3SG	ART.DEF.SG.OBL	bon	homme-SG.OBL	ART.DEF.PL.OBL	bon	ART.DEF.PL.OBL	Rom-PL

« Dieu il faut/ qu'il les brûle ceux-là. Qu'il garde la bonne personne/ les bons Roms. »

Ajia Varvara

Le corpus de récits collecté par Igla (1996) présente des cas intéressants de polydéfinitude. L'exemple (264) exemplifie le fait que le premier déterminant du syntagme puisse être un possessif (structure POSS N DEF A). L'exemple (265) illustre quant à lui que la préposition monosyllabique, même non fusionnée avec l'article défini, est rédupliquée dans un syntagme polydéfini.

(264) Romani (Igla 1996 : 269)

Mek-él-as	*m-i*	*kumnáta*	*i*	*bari*	*p-e*	*čha-én.*
laisser-3SG-IMPF	POSS.1SG-F.SG	belle_sœur(F.SG)	ART.DEF.F.SG	grand-F.SG	POSS.REFL-PL.OBL	enfant-PL.OBL

« Ma belle-soeur aînée laissait ses enfants. »

(265) Romani (Igla 1996 : 255)

D-el	*sovná*	*o*	*thagár*	*ekh*	*ðiatají,*	*d-el*	*p-o*
donner-3SG	ensuite	ART.DEF.M.SG	roi(M.SG)	un	ordre	donner-3SG	POSS.REFL-M.SG
thagaripé	*ka*	*o*	*rakló*	*ka*	*o*	*cəkn-ó*	
royaume(M.SG)	à	ART.DEF.M.SG	garçon(M.SG)	à	ART.DEF.M.SG	petit-M.SG	

« Ensuite le roi donne un ordre, il donne son royaume au petit garçon. »

La préposition est rédupliquée de façon systématique dans les appositions étroites nominales (PREP DEF N PREP DEF N) : *ka po dad ka o thagár*, à son père le roi, littéralement « à son père au roi » (Igla 1996 : 255), *ka o Jáni ka o Paríδi*, à Jani Paridi, littéralement « à Jani à Paridi » (Igla 1996 : 269), *ka mo phral ka o Jóryo*, à mon frère Yorgos, littéralement « à mon frère à Yorgos » (Igla 1996 : 269).

Parakalamos

Outre la structure polydéfinie à adjectif antéposé, *e parnes e grastes*, le cheval blanc (structure DEF A DEF N), (Matras 2004 : 76) note l'existence d'une structure à deux déterminants avec réplication du possessif : *mo baro mo phral*, mon grand frère (structure POSS A POSS N), où l'adjectif est antéposé. Cette réplication du possessif se présente également en l'absence d'adjectif : *me duj me phenja*, mes deux sœurs (POSS NUM POSS N) (Matras 2004 : 76). On peut questionner ici le statut du numéral, à considérer comme un déterminant ou un quantifieur. Dans ce cas précis, l'absence de flexion fait opter l'auteur pour un déterminant.

Lorsque le numéral porte la flexion nominale dupliquée du nom, l'auteur l'interprète comme un adjectif. C'est le cas pour un syntagme similaire au précédent, avec réduplication du possessif en

présence du numéral *mi dujendar mi phralendar*, avec mes deux frères, qui est analysé par (Matras 2004 : 75) non comme POSS NUM POSS N, mais comme POSS A POSS N. Le syntagme avec réduplication de l'article défini *e dujengoro e phralengoro*, des deux frères, est analysé non comme DEF NUM DEF N, mais comme DEF A DEF N. Enfin, le syntagme *olenza i trinenza i muršenza*, avec ces trois hommes, est analysé non comme DEM DEF NUM DEF N, mais comme DEM DEF A DEF N, avec trois et non quatre déterminants (Matras 2004 : 75). Toutefois, en tant qu'adjectif, le numéral devrait porter la flexion adjectivale et non la flexion nominale, surtout s'il est antéposé.

Komotini

La base de données en ligne Pangloss (Lacito 2001), créée par l'UMR 7107 Lacito, archive des corpus d'enregistrements en langues rares ou en danger. On y trouve plusieurs enregistrements en romani de Grèce (enregistrés par Evangelia Adamou) et du Mexique (par Evangelia Adamou et Cristian Pădure). Le conte *The louse and the Rom*, la puce et le Rom (Adamou 2006a), a été enregistré en 2006 en Thrace, auprès d'une petite communauté d'environ trois cents personnes vivant dans les banlieues de Komotini (Adamou 2010). Cette variété *vlax* est parlée par une petite minorité de Roms musulmans qui ont une longue tradition de trilinguisme romani-turc-grec. Leur variété a été influencée par le grec (langue nationale) mais surtout par le turc (langue locale et régionale) depuis l'époque de l'Empire ottoman (Adamou et Arvaniti 2011, Adamou 2010). Le conte présente une occurrence de syntagme polydéfini avec adjectif postposé :

(266) Romani (Adamou 2006a : ligne 74)

kaj	*lel*	*kotar*	*o*	*kaʃ*	*o*	*phabhardo*	*kaj*	*perel*	*palal*	*leste*
where	take.3SG	from	ART.SG.M	wood	ART.SG.M	burning	where	fall.3SG	behind	3SG.M.LOC

'He grabs a burning piece of wood and starts chasing him.'

On observe même une occurrence de syntagme polydéfini où un article défini « introduit » un démonstratif. Il s'agit d'un syntagme à trois déterminants (structure DEF DEM DEF N), le même que celui observé en Roumanie *o kodowa o ćhaw'*, ce garçon (litt. « le ce le garçon ») (cf. p. 416).

(267) Romani (Adamou 2006a : ligne 79)

avilo	*kaj*	*kova*	*than*	*kaj*	*ka*	*kerenas*	*marebava*
come.PRET.3SG.M	where	DEM.M	place	where	will	make.REM.3PL	war
bitʃalda	*le*	*o*	*gasavo*	*o*	*patiʃaj*.		
send.PRET.3SG	3SG.OBL.M	ART.SG.M	that.one	ART.SG.M	king		

'He arrives at that place where they were supposed to fight, where the king sent them.'

Toutefois la structure est très rare dans cette variété – ce que l'on peut attribuer à l'influence très importante du turc, qui est selon (Adamou 2016 : 159) sa principale langue de contact. En effet le turc non seulement ne comporte pas de polydéfinitude, mais il ne comporte pas non plus d'article défini (cf. section 1.4.1.2, p. 425).

1.3.5.3 Bulgarie

J'ai pu interviewer plusieurs Roms bulgares au cours d'une mission de terrain effectuée à Thessalonique à l'été 2014. Le locuteur de l'exemple suivant est un homme de vingt ans parlant une variété *vlax*. Il réside à Thessalonique depuis quatre ans mais est originaire de Teteven en Bulgarie. Voici un syntagme polydéfini avec adjectif antéposé issu d'une conversation quotidienne et prononcé avec un geste de monstration :

(268) *M20VLBg, Thessalonique, 26 août 2014*

Kava	*si*	*mir-o*	*amal*	*o*	*baro.*
DEM	être.3.PRES	POSS.1SG-M.SG	ami(M.SG)	ART.DEF.M.SG	grand-M.SG

« Celui-ci c'est mon meilleur ami (litt. mon grand ami). »

1.3.5.4 Hongrie

Le recueil de textes de Courthiade (2007) comprend un conte en romani de la variété hongroise ursari. Voici plusieurs phrases comprenant des syntagmes polydéfinis avec adjectif antéposé.

(269) Romani (Courthiade 2007)

O	*Thagár*	*o*	*Lol-ó*	*ésas*	*les*	*jekh*	*ćhăj*	*haj*
ART.DEF.M.SG	roi(M.SG)	ART.DEF.M.SG	rouge-M.SG	être.3.PRET	3.SG.OBL	un	fille(F.SG)	et
o	*Thagár*	*o*	*Parn-ó*	*sas*	*les*	*jekh*	*ćhăwó.*	
ART.DEF.M.SG	roi(M.SG)	ART.DEF.M.SG	blanc-M.SG	être.3.PRET	3.SG.OBL	un	fils(M.SG)	

« Le Roi Rouge avait une fille et le Roi Blanc avait un fils. »

(270) Romani (Courthiade 2007)

O	*Parn-ó*	*Thagár*	*gel-ó*	*te*	*ker-él*
ART.DEF.M.SG	blanc-M.SG	roi(M.SG)	aller.PRET-3SG.M	COMPL	faire-3SG
biáw	*le*	*Thagar-é-ça*	*le*	*Lol-é-ça.*	
mariage(M.SG)	ART.DEF.M.SG.OBL	roi-M.SG.OBL-INSTR	ART.DEF.M.SG.OBL	rouge-M.SG.OBL-INSTR	

« Le Roi Blanc alla organiser le mariage avec le Roi Rouge. »

(271) Romani (Courthiade 2007)

O	*grast*	*o*	*bar-o*	*ʒa-l*	*k-o*	*gonó*	*baró*	*haj*
ART.DEF.M.SG	cheval(M.SG)	ART.DEF.M.SG	grand-M.SG	aller-3SG	à-ART.DEF.M.SG	sac(M.SG)	grand-M.SG	et
o	*grast*	*o*	*cïgn-ó*	*ʒa-l*	*k-o*	*cïgn-ó*	*gonó.*	
ART.DEF.M.SG	cheval(M.SG)	ART.DEF.M.SG	petit-M.SG	aller-3SG	à-ART.DEF.M.SG	petit-M.SG	sac(M.SG)	

« Le grand cheval va au grand sac et le petit cheval va au petit sac. »

(272) Romani (Courthiade 2007)

Din-ă	*len*	*drom*	*k-o*	*Thagár*	*o*	*Parnó.*
donner.PRET-3SG	3.PL.OBL	chemin(MG)	à-ART.DEF.M.SG	roi(M.SG)	ART.DEF.M.SG	blanc-M.SG

« Il les mena au Roi Blanc. »

1.4 Scénario géographique

1.4.1. Dans les Balkans

1.4.1.1 Le romani

La Figure 76 représente les zones où nous avons relevé des syntagmes polydéfinis dans les Balkans. Les publications scientifiques (cf. section 1.3.1, p. 411) attestent de leur présence également hors de la région (Suède et Colombie) : nous supposons que c'est grâce à la migration des locuteurs *vlax*.

Figure 76 : Zone d'extension de la polydéfinitude dans les Balkans

Il semble que la majorité des variétés des Balkans connaisse la polydéfinitude et l'emploie à des fréquences diverses. Une telle extension rend difficile d'imaginer qu'il s'agisse chaque fois d'une innovation interne locale effectuée par des groupes dispersés : cela invalide notre **hypothèse A**. On préfère donc opter pour l'**hypothèse B**, l'emprunt à une autre langue.

Il convient de rappeler ici quelles sont les langues en contact avec les différents groupes de locuteurs. Nous savons que toutes les variétés de romani ont des L2 de contact ancienne, récente et actuelle (cf. section 2.2.1 du chapitre I, p. 34). Sur la base des publications historiques, de mes observations de terrain et des déclarations de mes informateurs, je fais l'hypothèse des L2 présentées en Tableau 7 (p. 43), répété ici :

Tableau 7: Histoire du contact de langues des variétés de romani d'Albanie

	Mečkar	Arli	Čergar I	Čergar II	Bamil	Kurtofi
L2 anciennes	albanais grec (slave du sud ?)	turc grec	langue romane (slave du sud)	turc grec (slave du sud)	grec (turc)	?
L2 récentes	albanais (grec)	grec albanais (turc)	slave du sud albanais (grec)	grec albanais (turc)	grec	?
L2 actuelles	albanais (grec) (italien)					

Qu'en est-il du grec, langue particulièrement importante pour le romani sur les plans diachronique et synchronique ? Il a une influence culturelle et religieuse très importante dans la région de Korçë – comme dans le reste de l'Albanie – même si sa connaissance effective varie aujourd'hui d'un locuteur à l'autre. Parmi les participants aux tests, six ont une connaissance active du grec parce qu'ils habitent ou ont habité en Grèce, se sont convertis au christianisme sur place et/ou travaillent ou ont travaillé avec des Grecs (en Grèce ou en Albanie). Six autres personnes ont une connaissance partielle de la langue en production et en compréhension, parce que leurs parents la leur ont transmise ou bien parce qu'ils ont travaillé avec des Grecs en Albanie. Vingt-deux personnes, notamment parmi les jeunes, disent ne pas connaître du tout le grec[1], mais leurs grands-parents oui. En ce qui concerne la religion, les Roms de Korçë que j'ai rencontrés sont orthodoxes de deuxième ou troisième génération. Les chrétiens de la première génération ont grandi en tant que musulmans puis se sont convertis à l'âge adulte ; ils ont maintenant 70 à 80 ans. Le reste des conversions a eu lieu durant les années 1990, à l'occasion de la guerre civile ou en raison du processus de migration (vers la Grèce ou l'Italie). Les Roms non-chrétiens que j'ai pu rencontrer sont soit laïcs, soit musulmans – auquel cas ils considèrent la conversion au christianisme comme une perte partielle de l'identité rom, une « gadžisation ». Parmi les locuteurs que j'ai pu rencontrer, ceux qui connaissent le grec le connaissent en raison de son prestige, de son utilité pour le travail, le commerce, la migration (ou les relations familiales avec les membres de la famille vivant en Grèce), plutôt qu'en raison de la religion.

1.4.1.2 Les autres langues

Le Tableau 99 présente quelles langues possèdent un article défini et/ou la polydéfinitude.

[1] Le grec, en tant que langue de prestige dans le domaine culturel et médiatique en Albanie, et en tant que langue de migration, est aujourd'hui en concurrence avec l'italien.

Tableau 99 : Présence de la mono- et de la polydéfinitude dans les autres langues des Balkans

	Article défini	Polydéfinitude
Albanais (tosque & guègue)	✓	(✓)
Roumain, moldave	✓	✓
Aroumain, mégléno-roumain	✓	✓
Grec	✓	✓
Bulgare	✓	×
Macédonien, serbe du sud-est	✓	×
Serbo-croate-bosniaque-monténégrin	×	×
Slovène	×	(✓)
Turc rouméliote	×	×
Hongrois	✓	×

1.4.1.2.1 Article défini

La quasi-totalité des langues des Balkans présente un article défini : c'est l'un des traits caractéristiques du *Sprachbund* (union linguistique) balkanique (Trubeckoj 1928, Sandfeld 1930, Banfi 1985, Assenova 2002, Tomić 2004, Sobolev 2004, Tomić 2006 : 1-48, Feuillet 2012) dans le domaine de la morpho-syntaxe (Rivero et Ralli 2001, Joseph 2001, Tomić 2006). Le romani n'est pas toujours mentionné dans les travaux portant sur cette aire, même ceux récents, tels que Feuillet (2012). Mais Courthiade (1986a, 1986b) ou encore Friedman (2000c) ont démontré qu'il en fait partie.[1]

En ce qui concerne l'article défini, on peut comparer le romani aux langues slaves. On sait que la plupart d'entre elles n'ont pas d'article défini… à l'exception de plusieurs langues slaves du sud : macédonien et bulgare (Mladenova 2007, Feuillet 2012), pomak (Adamou 2011), nashta (Adamou 2016 : 141), kajnas[2] (Courthiade 1988a), dialectes serbes du sud-est entre Prizren et le Timok (Tomić 2006 : 118-125), et même le slovène (Marušič et Žaucer 2014, Runić 2015).[3] L'article défini est donc un trait acquis ou en cours d'acquisition propre à cette aire linguistique.

Il en va de même pour le romani. En tant que langue indo-aryenne, il n'avait pas d'article défini avant d'entrer dans la région. Comme les langues slaves, il l'a créé par contact avec le grec et à partir de matériel linguistique hérité, le démonstratif (Boretzky 2000b, Matras 2002 : 280).[4] L'acquisition de l'article défini est un balkanisme majeur du romani : Matras (1994), Matras (1996) puis Matras

[1] (Tomić 2006 : 45-46) a ainsi pu citer le « romani balkanique » comme une langue faisant partie de l'union linguistique, et le mentionner tout au long de son ouvrage. Elle présente par exemple les articles défini et indéfini du « romani arli balkanique » (Tomić 2006 : 224-227). Toutefois, la source et l'auteur de ses exemples et explications ne sont pas mentionnés. Le texte en romani transcrit en p. 695 a été traduit du macédonien : nous pouvons faire l'hypothèse qu'il s'agit d'une variété d'arli macédonien.

[2] parler de Boboščica (Boboshticë) et de Drenovəne (Drenovë) au sud de Korçë en Albanie.

[3] A cette liste, nous pouvons ajouter le sorabe, langue parlée en Saxe, au Brandebourg et en Pologne, qui a acquis un article défini libre antéposé au nom, sur la base de son démonstratif, par contact avec l'allemand (Lötzsch 1996 cité par Matras et Sakel 2007 : 845).

[4] Un tel processus de grammaticalisation est courant (cf. section 1.3.3, p. 65, et Lyons 1999).

(2000a : 176) a qualifié le romani de « langue indienne balkanisée ». Cela signifie que la langue est indienne de par son vocabulaire et son fonctionnement grammatical originel, mais qu'elle a subi d'importants changements structuraux et typologiques (*extreme typological disruption*, Matras 2000a : 176). La plupart d'entre eux sont des traits communs aux autres langues des Balkans : Matras (2000a : 177) liste une série de onze traits balkaniques qui caractérisent la plupart des variétés de romani actuelles, même hors d'Europe. L'article défini en fait partie.

1.4.1.2.2 « Double » article défini

On constate, avec le Tableau 99, que la polydéfinitude a un caractère aréal. Elle correspond à des configurations grammaticales différentes. Les stratégies morpho-syntaxiques sont variables : les langues peuvent exprimer la définitude à l'aide d'un morphème libre ou d'un morphème lié. La réplication de la définitude peut se faire par la réplication (doubler l'article défini : grec standard et de Cappadoce, aroumain, slovène) ou par la combinaison d'un morphème libre et d'un morphème lié (marquage mixte : langues roumaines, albanais). La polydéfinitude peut être grammaticalement « obligatoire » (systématique : grec de Cappadoce, albanais) ou bien « facultative » (variable : grec standard, slovène, langues roumaines). Lorsque la polydéfinitude est obligatoire, elle est purement « grammaticale » et ne présente aucun apport sémantique particulier. Lorsqu'elle est optionnelle, on constate qu'elle présente au moins une fonction contrastive. Au vu de cette diversité, il est plus juste de parler de « marquage multiple » de la définitude plutôt que de « polydéfinitude ».

Lekakou et Szendrői (2012 : section 5) affirment que pour comporter la polydéfinitude, une langue doit également comporter un système de flexion nominale (dans la structure sous-jacente, une tête 'Kase' séparée) et la détermination des noms propres (dans la structure sous-jacente, une scission Def-D). Elles proposent une typologie des langues :

- Type 1 : déclinaison et déterminant obligatoire sur les noms propres (grec, bavarois)
- Type 2 : déclinaison mais pas de déterminant obligatoire sur les noms propres (allemand standard et septentrional)
- Type 3 : ni déclinaison ni déterminant obligatoire sur les noms propres (anglais, italien)
- Type 4 : pas de déclinaison, mais déterminant obligatoire sur les noms propres (catalan, variétés septentrionales de l'italien)

Seules les langues appartenant au type 1 peuvent comporter de la polydéfinitude. Kyriakaki (2011 : 3) contredit ce constat, en affirmant que l'albanais possède la flexion nominale et la détermination des noms propres, mais pas de contruction polydéfinie. Alexiadou (2014 : 119-120), quant à elle, valide ce constat – alors même qu'elle rejette l'analyse de l'article défini grec comme explétif.

Alexiadou (2014 : 111-121) propose une typologie tripartite de la détermination multiple : « *cases that involve interpretational effects in the presence of extra determiners, cases that lack such effects, and finally patterns associated with spurious articles. I labelled the first case the syntactic pattern of multiple realization, the second case, the morphological pattern, and the third case the spurious articles/determiners pattern* ». Le *pattern* syntaxique se distingue du *pattern* morphologique en ce

que le deuxième n'a aucun effet sur l'interprétation du syntagme, contrairement au premier. Trois critères d'analyse fondent cette tripartition : les effets d'interprétation liés aux adjectifs et aux déterminants, les restrictions sur les adjectifs, la sensibilité à la forme.

- Les langues relevant du *pattern* syntaxique sont le norvégien et le suédois (qui appartiennent au sous-type *split DP layer*) ainsi que le grec (qui appartient au sous-type *(in)definiteness spreading*). Le roumain est proche du grec en ce que l'article adjectival implique des propriétés adjectivales telles que l'interprétation restrictive, l'intersectivité, la prédicativité. Mais l'article grec a un effet pragmatique que n'a pas l'article roumain : le trait de familiarité.

- Les langues relevant du *pattern* morphologique sont l'hébreu (sous-type accord en définitude, « *spreading that is the result of copying of features* ») et l'albanais (avec des « *adjectival articles that are subject to conditions on insertion of vocabulary items* »). Les déterminants multiples hébreux impliquent une copie du trait de définitude, tandis que les déterminants multiples albanais sont sujets à des facteurs morphologiques et n'impliquent pas une simple copie de type accord. L'article défini albanais a fait selon elle l'objet d'une réanalyse.

- Les langues relevant du *pattern* du faux article sont le slovène et le chinois. Les constructions à déterminant multiple n'impliquent pas de déterminant à proprement parler, « *elements that realize a D° head, in agreement with Plank's implicational universal that no language will practice double (or multiple) articulation unless it also articulates its noun phrases singly* » Alexiadou (2014 : 112). Le slovène possède une structure de syntagme adjectival très riche.

Alexiadou (2014 : 118) explique la distribution des déterminants multiples dans les langues du monde par le principe suivant : « *Languages that have null D heads lack multiple realization of (in)definiteness in adjectival modification contexts that involve a relative clause structure* ». Ce principe implique deux conditions : que les adjectifs puissent apparaître comme prédicats dans des propositions relatives réduites et que le sujet de ces prédicats soit introduit par un déterminant nul. Ces deux conditions sont sujettes à variation dans les langues du monde. L'anglais, le français ou l'italien ont un déterminant nul qui peut rester silencieux lorsqu'il introduit des arguments en contexte de modification impliquant des propositions relatives. Ce n'est pas le cas du grec : c'est pourquoi ces langues n'ont pas de déterminants multiples.

Albanais

On peut consulter à ce sujet Campos (2009), Androutsopoulou (2001 : 162, 164, 179, 195) ainsi que Lyons (1999 : 71, 75-76), Plank (2003b : 358-359), Tomić (2006 : 186-203), Feuillet (2012) et Alexiadou (2014 : 84, 88-89).[1] L'albanais présente un système mixte de marquage de la définitude à l'aide d'un morphème lié postposé au nom (similaire à celui que l'on trouve en roumain, en suédois et en norvégien) et d'un morphème libre antéposé à l'adjectif. La structure albanaise n'est pas à proprement parler un cas de polydéfinitude, du moins pas sur le plan fonctionnel. Mais elle l'a probablement été dans le passé, et l'est en partie sur le plan formel.

[1] La similarité entre le grec et l'albanais a été notamment discutée par Androutsopoulou (2001), Campos (2009 : 1024-1027) et Alexiadou (2014 : 84-90).

L'article défini proprement dit est un morphème lié postposé en deuxième position du syntagme. Sa nature est discutée : article, désinence flexionnelle Androutsopoulou (2001 : 162)[1], suffixe ou clitique Lyons (1999 : 71, 75-76), enclitique de syntagme (qui s'attache au premier élément du syntagme, à l'exception d'un mot « faible » comme l'article adjectival). Lorsque le syntagme est introduit par un démonstratif, le nom peut prendre ou non cet article : le syntagme a alors la structure DEM N ou DEM N-DEF lorsque la référence est anaphorique.

L'autre marqueur est un morphème libre qui « accompagne » tout modifieur postnominal, adjectif, numéral ordinal ou syntagme génitival (Androutsopoulou 2001 : 182-186). Il doit lui être strictement adjacent. Notons que quelques adjectifs peuvent se passer de lui : il s'agit de ceux de la très minoritaire « classe 1 », tandis que ceux de la très majoritaire « classe 2 » le prennent obligatoirement (Androutsopoulou 2001 : 193). Cet article a la même origine que l'article défini proprement dit (Androutsopoulou 2001 : 164, 179, Lyons 1999 : 79-80), dont il a hérité la complexité et sa ressemblance morphologiques (Campos 2009, Lyons 1999 : 71, Alexiadou 2014 : 88-89). Il s'accorde en genre, nombre, cas et définitude avec le nom-tête (et non, par exemple, avec le nom du syntagme génitival) : c'est pourquoi Lyons (1999 : 79) parle de morphème d'accord associé à l'adjectif et Alexiadou (2014 : 84) de morphème polyvalent de définitude et de cas. Lyons (1999 : 76) va jusqu'à parler d'une sorte d'« auxiliaire » du syntagme nominal, une entité unique exprimant la définitude, le nombre et le cas. Ce morphème s'est toutefois spécialisé dans l'« accompagnement » du modifieur plutôt que dans la définitude : il apparaît aujourd'hui lorsque le syntagme est introduit par un article indéfini. Pour cette raison, beaucoup d'auteurs considèrent qu'il a perdu le trait défini. Il est toutefois morphologiquement sensible à la définitude. De plus, certains noms comme les noms de parenté, se comportent comme les adjectifs de classe 2 et doivent obligatoirement être précédés de cet article (en plus de l'article défini proprement dit, clitique postposé) (Plank 2003b : 380-381). Lorsqu'il porte sur de tels noms, l'article « adjectival » est alors bien un marqueur de définitude. Nous parlons d'article adjectival, car il accompagne l'adjectif non seulement en fonction épithète, mais également en fonction attribut du sujet. Sa nature est très discutée[2] ; il est considéré ou glosé :

- *déterminant adjectival* par Androutsopoulou (2001)
- *article adjectival* par Campos (2009)
- *article de connexion* et *article copulatif* par Feuillet (2012)
- *article* par Plank (2003b), Alexiadou, Haegeman et Stavrou (2007) et Alexiadou (2014)
- *the* ou *article* par Alexiadou (2014)
- *particule* par Friedman (2000a)

[1] Androutsopoulou (2001 : 195) discute s'il s'agit d'un affixe de mot, de syntagme ou d'un clitique.

[2] Campos (2009 : 1009) écrit : « *Albanian adjectival articles have received different names in the literature: "agreement clitics" in Tomic´ (2006), "preposed articles" in Trommer (2002), "adjectival determiners" in Androutsopoulou (2001), "adjective clitics" in Zymberi (2000), "attributive clitics" in Newmark (1997), "prepositive article" in Camaj (1984), "adjectival articles" in Newmark et al. (1982) and Dimitrova-Vulchanova and Giusti (1998), "connecting particles" in Mann (1997), "gelenkartikel" 'linking article' in Pani (2007), "pronombre conjuntivo" 'conjunctive pronoun' in Sanz Ledesma (1996), etc. The Akademia e Shkencave e Shqiperise¨ (Albanian Academy of the Sciences, 2002), in its Gramatika e gjuhe¨s shqipe (Grammar of the Albanian Language, GGJSH, henceforth) refers to them as nyja e mbiemrit (article of the adjective). Following Newmark's, Giusti's, Dimitrova-Vulchanova's as well as the Academy's terminology, I will refer to them as "adjectival articles".* »

- *particule adjectivale* par Lyons (1999)
- *the* par Franco, Manzini et Savoia (2013) qui le mettent sous « D » dans l'arbre syntaxique
- glosé morphologiquement par Friedman (1993)

Roumain

On peut consulter à ce sujet Campos (2005), Cornilescu et Nicolae (2012 : 1092), ainsi que Lyons (1999 : 75), Plank (2003b : 347-349), Campos et Stavrou (2004), Tomić (2006 : 125-152, 127-128), Cinque (2010 : 103), Feuillet (2012) et Alexiadou (2014 : 53-62).[1] Le roumain présente un système mixte de marquage de la définitude, à l'aide d'un morphème lié postposé au nom (similaire à celui que l'on trouve en albanais, en suédois et en norvégien) et d'un morphème libre antéposé à l'adjectif.

L'article défini proprement dit est un morphème lié postposé en deuxième position du syntagme. Il s'accorde en genre, nombre et cas avec le nom. Sa nature est discutée : clitique postposé (Lyons 1999 : 74-75), suffixe (Cornilescu et Nicolae 2012 : 1075-1076), enclitique de syntagme (en position de Wackernagel, qui s'attache au premier lexème) (Alexiadou 2014, Campos et Stavrou 2004 : 161-162).

L'autre marqueur est un morphème libre qui « accompagne » de manière optionnelle l'adjectif, le participe présent ou passé postposés (Cornilescu et Nicolae 2012 : 1088), ainsi qu'un syntagme prépositionnel complément du nom introduit par les prépositions *de*, de ou *cu*, avec (Plank 2003b : 348, Cornilescu et Nicolae 2012 : 1088). Etymologiquement, l'article adjectival provient du démonstratif distal mais ne s'interprète cependant pas de façon démonstrative ni déictique (Campos et Stavrou 2004 : 161). Sa flexion est différente du démonstratif : c'est celle d'un adjectif (Plank 2003b : 349). Lyons (1999 : 80-82), Campos et Stavrou (2004 : 161) et surtout Cornilescu et Nicolae (2012) l'étudient en détail. Il accompagne de manière obligatoire, lorsque l'adjectif est au superlatif (Lyons 1999 : 81). Il s'accorde en genre, nombre et cas avec le nom. Il s'emploierait seulement avec un adjectif pouvant se trouver en position prédicative (Cornilescu et Nicolae 2012 : 1092, Cinque 2010 : 103, 145), et il introduirait une proposition relative réduite sous-jacente (Cinque 2010 : 145). Etonnamment, la structure avec cet article n'est pas récursive : **băiatul <u>cel</u> înalt <u>cel</u> curajos, 'the courageous tall boy'*, le grand garçon courageux (Cinque 2010 : 145). Le terme d'article adjectival pourrait être justifié, comme en albanais, par le fait que cet article accompagne également l'adjectif en fonction attribut, ainsi que pour substantiver l'adjectif (*cel bun* serait sémantiquement équivalent à *bunul*, le bon (Plank 2003b : 349). Lorsque le numéral est ordinal (sauf « premier ») et antéposé au nom, il se combine avec l'article défini proprement dit : structure DEF ORD N avec marquage simple de la définitude. Lorsque le numéral ordinal est postposé, il se combine avec l'article adjectival : structure N=DEF ART ORD avec marquage double de la définitude (Lyons 1999 : 80). Cette alternative syntaxique rappelle celle du syntagme avec adjectif au superlatif. C'est pourquoi on peut considérer qu'il s'agit bel et bien d'un marqueur de définitude. Un numéral cardinal en fonction épithète dans un syntagme défini ne peut porter l'article défini : il ne peut porter que l'article adjectival (comme en albanais). Si le numéral cardinal a une fonction de déterminant quantifieur, il peut se passer de l'article adjectival : le syntagme n'est alors pas défini. Lorsque le modifieur est un adjectif possessif, un

[1] Pour une étude exhaustive sur la question de la détermination et du marquage simple ou double de la définitude en roumain, on pourra consulter Cornilescu et Nicolae (2012).

groupe génitival ou un numéral ordinal, et que le nom-tête présente déjà un autre modifieur, on emploie non l'article adjectival *cel* (issu du démonstratif distal *acel*), mais l'article adjectival *al* (apparenté à l'article défini *-ul*). Ce marqueur s'accorde, lui aussi, en genre et en nombre avec le nom-tête. Les deux articles *cel* et *al* s'emploient lorsque le possessif ou le numéral ordinal sont attributs du sujet. La nature de l'article adjectival est très discutée[1] ; il est considéré ou glosé :

- *article défini morphème libre* et *article adjectival* par Cornilescu et Nicolae (2012)
- *article de connexion* et *article copulatif* par Feuillet (2012)
- *article démonstratif (cel)* par la grammaire traditionnelle Feuillet (2012) et *article possessif (al)* Campos (2009 : 1031), Campos et Stavrou (2004 : 169)
- *article défini non-clitique* par Tomić (2006)
- *déterminant de forme libre* par Lyons (1999)
- *article démonstratif* par Alexiadou (2014)

Plank (2003b : 350) explique d'ailleurs qu'en **roumain ancien** (XIXème siècle et avant) et dialectal, on trouve des cas de triple marquage de la définitude, où l'adjectif portait lui aussi l'article défini clitique postposé : *soldat<u>ul</u> <u>cel</u> bun<u>ul</u>*, le bon soldat (structure N-DEF DEF A-DEF). La complexité d'un tel syntagme lui fait dire que la sous-partie adjectivale serait en fait nominalisée – ce pour quoi il n'apporte toutefois pas d'argument.

Aroumain

L'aroumain de Thessalie (Grèce) présente un système de marquage double de la définitude semblable à celui du roumain (Campos et Stavrou 2004, Campos 2005, Campos 2009 : 1026) : *stilo-<u>lu</u> <u>atsel</u> lai*, le stylo noir (structure N-DEF DEF A). Il connaît aussi un autre marquage double, consistant à réduplicuer l'article défini clitique postposé : *stilo-<u>lu</u> lai-<u>lu</u>*, le stylo noir (structure N-DEF A-DEF) (Tomić 2006 : 167-186, 171). Il connaît enfin un marquage triple, qui combine les deux stratégies précédentes : *stilo-<u>lu</u> <u>atsel</u> lai-<u>lu</u>*, le stylo noir (structure N-DEF DEF A-DEF) et qui apparaît lorsque le syntagme est « marqué contrastivement » (Tomić 2006 : 171). « *While in Modern Greek all adjectives to the immediate right of nouns are preceded by definite articles, in Aromanian, a demonstrative determiner has to intervene between an articled noun and a postnominal articled adjective. Articled adjectives occur to the immediate right of articled nouns only when they are in appositional relationship to the noun* » (Tomić 2006 : 171). Ainsi n'a-t-on pas « *stilolu lailu* », mais « *stilolu, lailu* », qui est une structure d'apposition lâche N=DEF, A=DEF. Notons que selon (Campos et Stavrou 2004 : 161), l'adjectif de la structure polydéfinie en grec et en l'aroumain présente la même propriété prédicative que l'adjectif postposé de l'italien.

L'article défini proprement dit est un morphème lié postposé en deuxième position du syntagme. Il s'accorde en genre, nombre et cas avec le nom – qui se fléchit également de son côté. Il accompagne le nom lorsque le syntagme présente un possessif, qui est alors postposé : N-DEF POSS. Il accompagne optionnellement le nom lorsque le syntagme présente un démonstratif et que celui-ci est postposé : DEM N ou bien N-DEF N. Contrairement à l'albanais, il ne peut être enclitique du second constituant.

[1] Sa nature est discutée de façon exhaustive par Cornilescu et Nicolae (2012).

L'autre marqueur est un morphème libre qui « accompagne » de manière optionnelle l'adjectif. Etymologiquement, il provient du démonstratif, auquel il est similaire, mais ne s'interprète cependant pas de façon démonstrative ni déictique : lorsqu'il est employé pour « accompagner » l'adjectif, *atsel* s'interprète bel et bien comme un article défini. Il est considéré ou glosé :

- *pseudo-article* par Campos (2005)
- *élément prédicatif* et *démonstratif* par Campos et Stavrou (2004)
- *déterminant démonstratif* par Tomić (2006)

Mégléno-roumain

Le mégléno-roumain de Macédoine présente les caractéristiques de l'aroumain (Tomić 2006 : 152-167, 155-156), avec un double marquage de la définitude : article défini proprement dit (clitique du nom) et l'article adjectival (antéposé à l'adjectif). Comme l'adjectif est antéposé en mégléno-roumain, on a une structure DEM A=DEF N. Si l'adjectif est « hautement emphatisé », l'adjectif mégléno-roumain est postposé au nom et porte lui aussi l'article défini clitique postposé : il s'agit d'une apposition lâche N-DEF, A-DEF selon (Tomić 2006 : 155). L'article adjectival, *ţel*, provient du démonstratif et est identique à lui.

Grec de Cappadoce

Le grec de Cappadoce est une variété de grec en contact séculaire avec le turc et qui présente un article défini libre semblable à celui du grec standard. Toutefois, en présence d'un adjectif épithète, la polydéfinitude est systématique (Karatsareas 2011, Karatsareas et Lekakou 2014, Lekakou et Karatsareas 2016), à quelques exceptions près (nominatif des anciens noms masculin et féminin, Karatsareas 2013). L'adjectif est obligatoirement antéposé : DEF A DEF N, *DEF N DEF A. Selon Karatsareas et Lekakou (2014) et Lekakou et Karatsareas (2016), il s'agit d'un accord en définitude sans aucun effet sémantique particulier, où les traits de définitude et de nombre sont simplement « copiés » du nom à l'adjectif. Ce phénomène rappelle l'accord en définitude des langues sémitiques (Rubin 2005b, Pat-El 2009)[1].

Bulgare

Le bulgare a un article défini clitique postposé (Avgustinova 1998, Feuillet 2012, Lyons 1999 : 73-74, Tomić 2006 : 86-108), mais ne présente pas de polydéfinitude.

[1] … où l'adjectif est toutefois postposé. Les langues sémitiques possèdent un article défini clitiqué antéposé au nom, qui est rédupliqué sur l'adjectif épithète. On peut dire que l'adjectif s'accorde en genre, nombre, cas et définitude avec le nom (Lyons 1999). Corbett (2006 : 135-137) discute toutefois cette notion d'accord en définitude.

Macédonien

Le macédonien a une série d'articles définis clitiques postposés (Foulon 1997, Feuillet 2012, Tomić 2006 : 49-86, Lyons 1999 : 73-74), mais ne présente pas de polydéfinitude.

Serbo-croate

Le serbo-croate n'a pas d'article défini (Despić 2013, Tomić 2006 : 108-118, Feuillet 2012, Thomas et Osipov 2012, Cinque 2010 : 99-102). La définitude se marque sur la flexion adjectivale « longue » ou « courte » et la prosodie (ton, place de l'accent). Il s'agit d'un héritage du système vieux-slave (Lyons 1999 : 82-83).

Les dialectes serbes du sud-est, entre Prizren et le Timok, ont un système semblable à celui du macédonien (Tomić 2006 : 118-125), avec articles définis mais sans polydéfinitude.

Slovène

Le slovène ne présente pas, en principe, d'article défini. La définitude se marque sur la flexion adjectivale et surtout sur la prosodie (ton, place de l'accent), héritage du système vieux-slave (Lyons 1999 : 82-83).

Les recherches sur le slovène non-standard ont toutefois montré l'émergence d'un article défini, *ta* (Trovesi 2004). Il s'agit d'une forme phonologiquement réduite du pronom démonstratif grammaticalisé, et qui n'apparaît qu'en présence d'un adjectif (antéposé) dans le syntagme nominal, aussi bien défini qu'indéfini (Marušič et Žaucer 2006, Marušič et Žaucer 2008, Marušič et Žaucer 2014). Le syntagme présente une structure originale à deux articles successifs : DEF DEF A N. Le phénomène est itératif : DEF DEF A DEF A N s'il y a deux adjectifs. Lorsque l'adjectif est dérivé ou au superlatif, on peut même trouver DEF A-DEF-A N, où le déterminant vient scinder l'adjectif (Marušič et Žaucer 2008 : 105). Selon (Marušič et Žaucer 2008 : 115), l'adjectif ainsi articulé est dérivé d'une proposition relative réduite et doit être prédicatif. Cette polydéfinitude relève de la variation diaphasique (et probablement diastratique), mais aussi de la variation diatopique. En effet, Runić (2012) et Runić (2015) a observé la même grammaticalisation du démonstratif *te* en article défini dans le dialecte slovène de Resia.

Marušič et Žaucer (2008 : section 3) étudient les différences entre cette structure slovène et la polydéfinitude en suédois et en grec. Ils soulignent les similitudes avec la « polydéfinitude » en albanais et en chinois, ainsi qu'avec la forme longue des adjectifs en slovène standard et en serbo-croate. Selon Alexiadou (2014 : 90-92), l'article défini slovène n'est pas un vrai déterminant qui réaliserait une tête D°. Surtout, il violerait l'implication universelle qui veut qu'une langue ne puisse pratiquer l'articulation multiple en présence d'adjectifs, si elle ne pratique pas déjà l'articulation simple en l'absence d'adjectif : DEF A N n'existe pas en slovène.

Turc

Le turc standard n'a pas d'article défini (Mürvet 1991 : 9, Lyons 1999 : 50,96), et le turc rouméliote non plus (Friedman 2003a, Friedman 2006).

Hongrois

Le hongrois a un article défini (Moravcsik 2003, Plank 2003b : 337-338, 387 et Szendrői 2010) mais ne présente pas de polydéfinitude.

1.4.1.2.3 Quelle langue d'emprunt ?

A quelle langue le romani albanais a-t-il emprunté la polydéfinitude ? Il pourrait s'agir de l'albanais, langue principale de contact et qui présente une structure proche. Une telle hypothèse ne permet toutefois pas d'expliquer pourquoi on trouve cette structure en romani d'autres pays (Roumanie, Moldavie, Grèce, ex-Yougoslavie…). Il faut alors postuler chaque fois un emprunt unilatéral à la langue de contact dominante locale : cette hypothèse est peu « économique ». Et comment expliquer que des variétés de romani puissent présenter de la polydéfinitude alors que la langue de contact dominante (serbo-croate, bulgare, macédonien) n'en présente pas ? De plus, cette hypothèse nous semble peu plausible en raison des « types » et usages hétérogènes de la polydéfinitude dans les langues balkaniques. La polydéfinitude semble au contraire homogène dans toutes les variétés de romani concernées.

C'est pourquoi l'emprunt à une autre langue nous semble plus probable – et à une langue unique. Cette langue doit bénéficier d'une grande influence et d'une vaste extension. Il s'agit plus probablement du grec, qui a tant influencé le vieux-romani puis les variétés balkaniques du romani, sur les plans morphologique, syntaxique, lexical et typologique. On sait que c'est par convergence avec le grec que le romani a acquis son article défini. Cela expliquerait pourquoi, par exemple, les variétés de Roumanie, qui ne sont plus en contact avec le grec aujourd'hui, puissent continuer à utiliser la structure polydéfinie. En Albanie, le grec est une langue de prestige sur les plans culturel et religieux, mais aussi la langue du pays voisin. En outre, la polydéfinitude du grec est la plus proche de la polydéfinitude du romani (cf. section 2.2.1, p. 147) : réplication de l'article défini, qui est un morphème libre, absence de poly*in*définitude, contextes d'emploi similaire… La polydéfinitude des autres langues balkaniques présente soit des différences au niveau formel (absence d'un double article défini, ou absence de réplication formelle) ou fonctionnel (polydéfinitude obligatoire, contextes d'emploi différents…). Nous en déduisons que c'est au grec que le romani a emprunté, de très bonne heure, la polydéfinitude. Le contact avec une langue locale présentant une forme similaire a pu favoriser le maintien de cette structure : c'est le cas de l'albanais en romani d'Albanie.

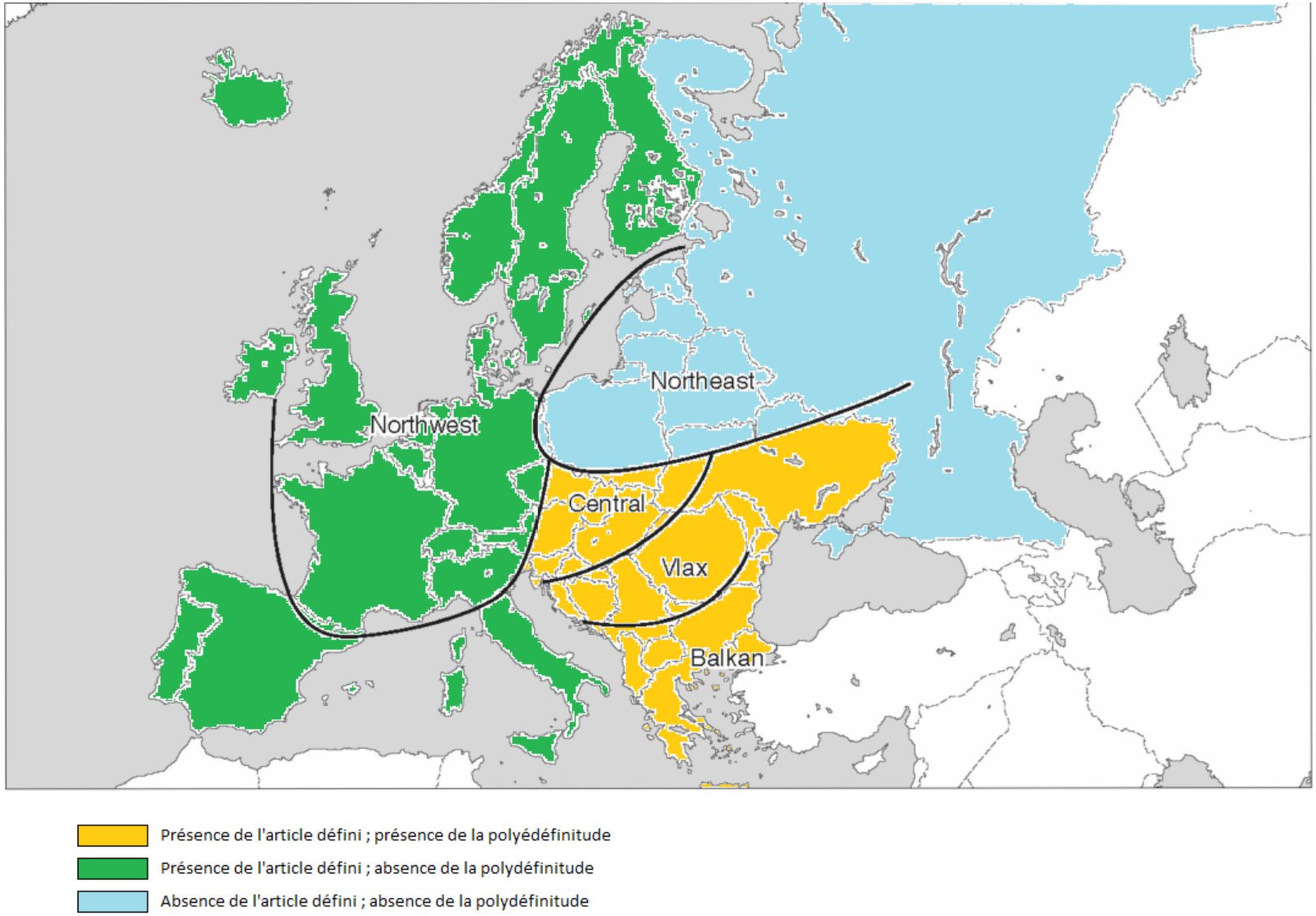


Figure 77 : Aires des isoglosses de la mono- et de la polydéfinitude en Europe

1.4.2. Hors des Balkans

1.4.2.1 Le romani

En considérant la classification dialectale de la Figure 1, p. 33, on peut émettre l'hypothèse selon laquelle les zones macro-dialectales générales correspondraient aux isoglosses de la monodéfinitude (présence de l'article défini) et de la polydéfinitude (possibilité de répliquer l'article). Le Tableau 100 indique quelles aires macro-dialectales sont dotées de l'article défini et de polydéfinitude. La Figure 77, modification de la carte de Matras (2010 : 56), permet de localiser ces aires.

Tableau 100 : Hypothèse de présence de la mono- et de la polydéfinitude en Europe

	Article défini	**Polydéfinitude**
Balkan	✓	✓
Vlax	✓	✓
Centre	✓	✓
Nord-est	×	×
Nord-ouest + Italie + Espagne	✓	×

Tableau 101 : Présence de la mono- et de la polydéfinitude dans les principales langues d'Europe

Zone	Langue	Article défini	Polydéfinitude
Est	Russe	×	×
	Ukrainien	×	×
	Biélorusse	×	×
	Polonais	×	×
	Tchèque	×	×
	Slovaque	×	×
Nord	Letton	×	×
	Lituanien	×	×
	Estonien	×	×
	Finlandais	×	×
	Suédois	✓	✓
	Norvégien	✓	✓
	Féroïen	✓	✓/×
	Islandais	✓	×
Ouest	Danois	✓	×
	Frison	✓	×
	Allemand	✓	×
	Bavarois	✓	✓
	Yiddish	✓	✓

Zone	Langue	Article défini	Polydéfinitude
Ouest	Néerlandais	✓	×
	Flamand	✓	×
	Anglais	✓	×
	Gaélique d'Irlande, d'Ecosse	✓	×
	Gallois	✓	×
	Breton	✓	×
	Français	✓	×
Sud	Basque	✓	×
	Catalan	✓	×
	Espagnol	✓	×
	Galicien	✓	×
	Portugais	✓	×
	Occitan	✓	×
	Franco-provençal	✓	×
	Italien, corse	✓	×
	Sarde, sicilien	✓	×
	Maltais	✓	✓
	Chypriote	✓	✓
	Turc	×	×

1.4.2.2 Les autres langues

Le Tableau 101 présente quelles langues possèdent un article défini (en tant que morphème libre ou lié) et/ou la polydéfinitude.[1]

On constate que beaucoup de langues ne présentent pas d'article défini, notamment dans l'aire balte et slave de l'Europe du nord-est. Cela coïncide avec l'absence d'article défini en romani des mêmes régions (cf. section 1.3.1.1 du chapitre I, p. 107).

On constate que la polydéfinitude est rare hors des Balkans. On peut faire l'hypothèse que cela coïncide avec l'absence de polydéfinitude en romani des mêmes régions. On la trouve essentiellement en Scandinavie (langues germaniques), à Chypre (où l'on parle grec) et à Malte. Elle présente un caractère aréal dans les Balkans et l'Europe centrale, ainsi qu'un caractère génétique en ce qui concerne langues sémitiques et le groupe des langues germaniques (septentrionales et haut-allemandes). Les études montrent que la répétition de l'article indéfini est rare, plus encore que la répétition de l'article défini.

Les ouvrages de Plank (2003b) et Alexiadou (2014) présentent une analyse inter-linguistique de la double articulation dans de nombreuses langues d'Europe et au-delà. Tous deux en déduisent qu'il n'existe pas d'explication unitaire, pas d'analyse unique qui fonctionnerait pour tous les *patterns* de marquage multiple de la définitude dans les langues du monde.

Langues baltiques

En **lituanien** et en **letton**, la définitude se marque sur le paradigme flexionnel des adjectifs. Lyons (1999 : 82-85) traite la question de la flexion adjectivale définie dans les langues baltes (lituanien, letton). Ce paradigme provient d'un ancien pronom affixé. On peut consulter Penanros (2015) pour le lituanien, Lyons (1999 : 83-84) pour le lituanien et le letton, Menantaud (2015) et Menantaud (2011) pour le letton.

Langues slaves

On peut consulter (Zlatić 2014) et (Rappaport 2014) pour les langues slaves en général. (Cinque 2010 : 108-110) examine en particulier les paradigmes d'adjectifs à forme courte et longue en **russe**. (Trovesi 2004) examine la génèse d'article défini en **tchèque** et en **sorabe**.

[1] Nous ne tenons pas compte, ici, du cas particulier de l'adjectif au superlatif. Il provoque une structure similaire à la polydéfinitude en français (Alexiadou 2014 : 68-75, Plank 2003b : 360-363), en italien et en romanche (Plank 2003b : 361), ainsi qu'en maltais (Plank 2003b : 361-362).

Finnois

Le livre de Laury (1997) montre qu'un article défini est en cours de grammaticalisation en finnois oral, à partir du démonstratif *se*.

Langues germaniques

Schoorlemmer (2012) dresse une étude comparative du marquage de la définitude dans les langues germaniques. Il distingue d'une part le suédois, le norvégien et le féroïen, qui présentent un marquage double, d'autre part l'anglais, le néerlandais et l'allemand, qui présentent un marquage simple. Le danois et l'islandais présentent également un marquage simple, de type différent. De plus, l'adjectif (antéposé) présente des marques de flexion « forte » (indéfinie) ou « faible » (définie), comparable à celui d'autres langues germaniques comme l'allemand (Leu 2009, Dahl 2004 : 150-151, Lyons 1999 : 82-85). Pour les langues scandinaves à double marquage, on peut également consulter Dahl (2004), Marit (2005) et Alexiadou (2014 : 62-68).

Alexiadou (2014 : 113-114) expose la différence entre les différentes langues scandinaves concernant la définitude et la place de l'adjectif. Elle propose un scénario structurel diachronique afin d'expliquer la variation d'une langue à l'autre, notamment la rupture entre islandais et danois d'une part, norvégien et suédois d'autre part. Les Tableau 102 et Tableau 103, issus de Schoorlemmer (2012 : 110, 153) résument les différents types de marquage dans les langues germaniques.

Tableau 102 : Le marquage de la définitude dans les langues germaniques (Schoorlemmer 2012 : 110)

PATTERNS OF DEFINITENESS MARKING IN GERMANIC

Pattern	Adj. absent	Adj. present	Languages
1	DEF N	DEF A N	German, Dutch, English
2	N-DEF	DEF A N-DEF	Swedish, Norwegian, Faroese
3	N-DEF	DEF A N	Danish
4	N-DEF	A N-DEF	(modern) Icelandic

Tableau 103 : Variation morphologique de la définitude dans les langues germaniques (Schoorlemmer 2012 : 153)

CROSS-LINGUISTIC VARIATION IN DEFINITENESS MARKING CAUSED BY MORPHOLOGICAL DIFFERENCES

	Freestanding + suffix	Freestanding only	Suffix only
Local Dislocation before Chain Reduction	N-DEF, DEF A N-DEF (Swedish, Norwegian, Faroese)	DEF N, DEF A N (German, Dutch)	N-DEF, A N-DEF (Icelandic)
Chain Reduction before Local Dislocation	N-DEF, DEF A N (Danish)		

Langues germaniques septentrionales

Le **suédois** présente deux articles définis différents : un article défini lié, suffixe ou clitique postposé au nom (qui apparaît en l'absence d'adjectif) et un article défini libre antéposé au nom (qui n'apparaît qu'en présence d'un adjectif) (Börjars 1994). L'adjectif est lui-même décliné avec la flexion « faible » définie (Plank 2003b : 354-355, Lyons 1999 : 78, 82-85). En la présence d'un adjectif, les deux articles définis doivent apparaître simultanément. En la présence de plusieurs adjectifs, la structure n'est toutefois pas récursive. Alexiadou (2014 : 95-96, 102-104) explique que les dialectes suédois du nord présentent une réduplication de l'article indéfini itérative : structure IND N IND A (IND A). Pour la variation dialectale du suédois, on pourra consulter Dahl (2004 : 159-176.).

Le **norvégien nynorsk** pratique le même système de définitude que le suédois, avec un article libre antéposé et article clitique postposé au nom (Dahl 2004). Alexiadou (2014 : 102-103) explique que les dialectes norvégiens du nord présentent une réduplication de l'article indéfini itérative : structure IND N IND A (IND A).

Le **norvégien bokmål** et le **danois** présentent eux aussi deux articles définis. En l'absence d'adjectif, c'est le clitique postposé au nom qui apparaît. En présence d'un adjectif ou d'un numéral, c'est l'article libre antéposé au nom (et à l'adjectif). De cette façon, le syntagme nominal présente toujours un seul article défini. De son côté, l'adjectif en syntagme défini se fléchit sur le paradigme faible (Lyons 1999 : 77, Dahl 2004 : 155).

Le **féroïen** oral présente un marquage double semblable au suédois, tandis que le féroïen écrit présente un marquage simple semblable au danois, avec article libre antéposé au nom ou article clitique postposé au nom… Il peut également ne présenter aucun marquage, laissant la seule flexion faible de l'adjectif indiquer que le syntagme est défini (Plank 2003b : 357).

L'**islandais**, comme le danois, connaît l'article libre antéposé et l'article clitique postposé au nom (qu'il n'emploie jamais en même temps), ainsi que la flexion faible de l'adjectif. Il ne connaît donc que le marquage simple de la définitude (Plank 2003b: 357). Selon Plank (2003b : 388), le **vieil-islandais** présentait une variation entre marquage simple à la danoise (au moyen d'un article ou de l'autre) et marquage double à la suédoiss (au moyen des deux articles combinés). Dahl (2004 : 150) cite même un exemple de marquage triple au moyen de deux article libres et de l'article lié.

Langues germaniques occidentales

Pour l'**anglais**, on peut consulter Master (1997) et Lyons (1999). Alexiadou (2014 : 3) note que la réplication de l'article indéfini existe dans certains dialectes de l'anglais : *if he was in a such a desperate hurry*, s'il était si désespérément pressé.

Pour l'**allemand**, on peut consulter Plank (2003b : 340-341, 366-367), Leu (2009) et Cinque (2010 : 111). L'allemand présente un article défini qui est un morphème libre antéposé au nom. Comme dans les langues slaves et baltiques, la définitude se marque également sur la flexion adjectivale : l'adjectif se décline selon un paradigme différent (dit « faible ») de celui des syntagmes indéfinis (dit « fort »). La scission entre les paradigmes dépasse les deux articles défini et indéfini, puisque tous les déterminants se classent d'un côté ou de l'autre.

Selon Plank (2003b : 340-341), l'allemand familier présente une structure polydéfinie optionnelle et peu fréquente. L'adjectif utilisé est postposé. Il revêt un caractère de commentaire, une connotation négative et apporte une information non prédictible : *Vergiß* <u>*das*</u> *Schnitzel(,)* <u>*das*</u> *zähe!* « Laisse tomber l'escalope trop cuite ! ». Il ne s'agit pas d'une apposition, car la partie « adjectivale » ne porte pas d'accent secondaire et n'est pas séparée de la partie « nominale » par une pause. Elle manque selon lui d'indépendance syntaxique et ne constitue pas un SN plein, malgré son article défini.

Selon Plank (2003b : 374-375), le **bavarois** réplique lui aussi l'article défini dans le même genre de contextes (voir aussi Lekakou et Szendrői 2012 : 143-144, contredites par Alexiadou 2014 : 120). Le bavarois va plus loin, puisque la polydéfinitude est possible en l'absence d'adjectif, avec deux articles définis encadrant alors le nom[1]. L'auteur précise que la forme de l'article défini est alors la forme forte, proche du démonstratif. Lorsque le syntagme ne comporte pas d'adjectif, la forme faible est cependant acceptable. La polydéfinitude est un cas d'« emphase spéciale » employée des phrases exclamatives ou interrogatives, surtout par les locuteurs masculins (Plank 2003b : 374). Selon l'auteur, elle confère aux propos « une saveur émotionnelle caractéristique ». C'est pourquoi les mots employés doivent avoir un potentiel sémantique affectif.

Lekakou et Szendrői (2012 : 143-144) soulignent le parallèle entre le bavarois et le grec : dans les deux langues, l'article défini s'accorde morphologiquement avec le nom, est obligatoire avec les noms propres, la polydéfinitude semble similaire. Cette comparaison est fermement critiquée en note par Alexiadou (2014 : 120), qui signale que la polydéfinitude est possible avec adjectif antéposé (DEF A DEF N) en grec, mais non en bavarois. Toutefois selon Plank (2003b : 374-375), locuteur de la langue, le bavarois a une syntaxe très souple dans les contextes d'emphase affective : *die Sau die*, le gros porc (DEF N DEF) avec double article et sans adjectif, *die bsuffane Sau die*, le gros porc bourré (DEF A N DEF) avec double article et adjectif antéposé non articulé, *die Sau die bsuffane* (DEF N DEF A) avec double article et adjectif postposé article, *die bsuffane Sau die bsuffane* (DEF A N DEF A) avec double article et double adjectif antéposé et postposé articulé.

Le bavarois peut également répliquer l'article indéfini en cas d'emphase et d'intensification du sens de l'adjectif (Plank 2003b : 366-367, 375, Alexiadou 2014 : 3, 98-100). Elle n'est possible qu'en présence d'un adjectif. Selon (Plank (2003b), la partie « adjectivale » du syntagme ne constitue pas un syntagme nominal indépendant, malgré la présence d'un deuxième article. Si on la considère comme un syntagme nominal indépendant, du moins les liens entre celui-ci et la partie « nominale » sont-ils beaucoup plus étroits que dans un *afterthought* ou dans une apposition ordinaire (Plank 2003b : 375).

[1] Ce phénomène n'a lieu que lorsque la tête est un nom, jamais un pronom.

Selon Leu (2009 : 297), le **suisse alémanique** réplique l'article indéfini (voir aussi Alexiadou 2014 : 95-97).

Le **yiddish** présente une structure polydéfinie selon Taube (2012 : 467-468), Jacobs (2005 : 173-175) Plank (2003b : 345-346) et Dahl (2004 : 149). Cela n'est pas étonnant, puisque le yiddish est une variété germanique appartenant à la branche du haut-allemand (*Oberdeutsch*), comme l'alémanique (alsacien, souabe, suisse alémanique) et le bavarois (de Bavière, mais aussi l'autrichien). L'article défini du yiddish est un morphème libre antéposé au nom et à l'adjectif. La structure canonique est DEF A N avec adjectif antéposé (Taube 2012 : 467). La structure optionnelle est DEF N DEF A avec réplication de l'article défini et adjectif (ou numéral) postposé. Comme en grec, l'adjectif est antéposé et ne peut être postposé qu'à la condition d'être accompagné d'un article défini (Taube 2012 : 467). Comme en bavarois, la réplication de l'article indéfini est aussi possible (Taube 2012 : 467).

En syntagme polydéfini ou polyindéfini, la morphologie de l'adjectif postposé est complexe : il s'accorde en genre, nombre et cas avec le nom lorsque ce dernier est féminin ou masculin, mais de façon variable lorsqu'il est neutre[1]. Elle trahit un « relâchement » du lien entre l'adjectif et le nom – sans que l'on puisse aller jusqu'à dire qu'il s'agit d'un cas d'apposition ou d'*afterthought* selon (Plank 2003b : 346). Jacobs (2005 : 242) et Taube (2012 : 467), quant à eux, vont jusqu'à considérer la partie « adjectivale » de la structure DEF N DEF A comme un « SN séparé en apposition ».[2]

Une rapide étude que nous avons menée sur un ensemble de chansons et de contes fait apparaître que la structure est en réalité plus usitée avec l'article indéfini qu'avec l'article défini. Notre hypothèse est que la proximité phonétique entre l'article **indéfini yiddish** אַ *a*, et l'article **défini hébreu** ה *ha* contribue à cette préférence. L'adjectif hébreu étant postposé et porteur d'un article défini en raison de l'accord en définitude, la structure canonique DEF-N DEF-A aurait pu servir de modèle et exercer une attraction sur le yiddish. Les locuteurs favoriseraient ainsi DEF N DEF A sur DEF A N, et surtout IND N IND A (homophone de l'hébreu) sur IND A N. Reste à expliquer quelles raisons amènent un locuteur à utiliser une construction complexe avec adjectif postposé et double article (défini ou indéfini), plutôt qu'une construction simple avec adjectif antéposé et article unique.

Langues celtiques

Toutes les langues celtiques sont dotées d'un article défini, le breton disposant en outre d'un article indéfini (Borsley et Roberts 1996 : 38) (sur l'irlandais, voir aussi Ball et Müller 2010 : 188). Mais elles ne présentent pas de polydéfinitude.

[1] Si le genre grammatical est distinct du genre référentiel (*meydl*, fille, est un nom neutre), l'accord « logique » (féminin) est tout autant possible que l'accord « grammatical » (neutre). L'accord grammatical en question n'est pas celui de l'adjectif épithète mais celui de l'adjectif attribut, substantivé et introduit par un article indéfini.

[2] Mais Jacobs (2005 : 242) considère tout modifieur postnominal comme un « SN séparé propre » et une apposition, même le possessif : *der brudər majnər*, mon frère, « le frère mien ».

Basque et sarde

Pour le basque, on peut consulter Etxeberria et Giannakidou (2014) et Plank (2003b : 339-340). Pour le sarde, on peut consulter Bentley (2004).

Maltais

L'article défini du maltais est un morphème lié antéposé au nom, que Fabri (2001 : 151-157) qualifie d'article clitique. Lorsqu'un adjectif modifie le nom, il peut être accompagné lui aussi par l'article défini. Le maltais présente donc deux structures possibles, DEF-N A avec marquage simple ou DEF-N DEF-A avec marquage double de la définitude. On peut consulter à ce sujet Plank et Moravcsik (1996), Fabri (2001), Plank (2003b : 350-351, 382) et Cinque (2010 : 98-99). De quoi dépend cette variation ?

Selon Corbett (2006 : 136), le marquage double DEF-N DEF-A donne à l'adjectif un sens contrastif. Selon Plank (2003b : 350), il s'emploie si l'adjectif est restrictif et aide à mieux identifier le référent du nom. Selon Fabri (2001 : 161), il s'emploie seulement avec certains types de nom et de référence, plus exactement s'il y a un ou plusieurs référents possibles qui peuvent être dénotés par le nom dans le contexte. S'appuyant sur Löbner (1985), il distingue les noms « sortaux » des noms « relationnels », qui comprennent à leur tour les noms « fonctionnels ». Les noms sortaux ont une référence ambiguë car le nombre de référents possibles dans le contexte d'énonciation est ouvert. Ils sont donc employés de préférence avec des adjectifs articulés (Fabri 2001 : 162-163). La présence du marqueur de définitude sur l'adjectif indique que le locuteur veut choisir un individu particulier au milieu d'autres individus similaires dans le contexte. L'utilisation du marqueur de définitude dépend donc de l'ambiguïté référentielle du nom dans le contexte d'énonciation, c'est-à-dire la conscience qu'il pourrait y avoir un éventuel choix entre plusieurs référents, donc de la connaissance des référents potentiels dénotés par le nom.[1] On voit que la variation n'implique pas la grammaticalité ou l'agrammaticalité, mais la pertinence pragmatique. Un tel usage de la polydéfinitude est similaire à celui que l'on trouve en romani d'Albanie et en grec, tant dans sa forme que dans son usage.

La variation dépend aussi du type d'adjectif (Fabri 2001). L'absence de marqueur de définitude sur un adjectif non restrictif est agrammaticale en maltais. Un adjectif au sens intrinsèquement restrictif n'a, quant à lui, pas besoin d'être articulé parce qu'il écarte toute possibilité d'un choix référentiel (Fabri 2001 : 168).[2] Cinque (2010 : 98) parle d'adjectifs non-intersectifs. D'autres classes d'adjectifs ne peuvent s'articuler : les numéraux ordinaux, les adjectifs classificatoires (Cinque 2010 : 98, Plank et Moravcsik 1996 : 188), les adjectifs de nationalité issus de noms propres (Plank et Moravcsik 1996 : 188, Plank 2003b : 351, Fabri 2001 : 158).

Cette variation amène Fabri (2001 : 159, 170) à nier que la polydéfinitude maltaise relève d'un phénomène d'accord, comme dans les autres langues sémitiques. Elle est déterminée par des facteurs

[1] Un adjectif comme « chaud » n'est donc pas articulé s'il porte sur le soleil, puisque l'on sait, grâce à notre expérience du monde, qu'il n'existe qu'un seul soleil dans notre galaxie et qu'il est intrinsèquement chaud (Fabri 2001 : 164, repris par Cinque 2010 : 98). La référence est non-ambiguë parce qu'il n'y a pas de choix possible à effectuer entre plusieurs soleils chauds – à moins de se trouver dans un contexte de science-fiction.

[2] Ainsi *ewlieni*, unique, *aħħari*, ultime, *waħdani*, seul, *principali*, principal ou *primarju*, primaire, ne peuvent-ils jamais porter de l'article adjectival car ils font référence à une entité unique et excluent un choix (Fabri 2001 : 169).

sémantiques et pragmatiques. Du point de vue formel, l'accord en définitude ne devrait souffrir d'exception, l'adjectif devrait toujours covarier avec le nom en définitude (Fabri 2001 : 160). Le marqueur de définitude n'a selon lui aucun statut syntaxique. L'adjectif « articulé » ne constitue pas un syntagme nominal propre et autonome, en d'autres termes, la présence de l'article ne projette pas de DP propre, car l'article n'occupe pas une position de déterminant dans la structure sous-jacente du syntagme nominal. Plank (2003b : 351) précise qu'il n'existe pas de différence structurelle entre les structures DEF-N DEF-A et DEF-N A ; les adjectifs articulés ne sont pas moins liés au nom, ne lui sont pas structurellement plus distants, que les adjectifs non articulés. Leur degré de nominalité n'est pas non plus supérieur.

Turc

Le turc n'a pas d'article défini (Mürvet 1991 : 9, Lyons 1999 : 50,96). Il présente en revanche un article indéfini grammaticalisé depuis le numéral *bir*, un. Plank (2003b : 351-352) note qu'en syntagme indéfini, en cas d'« emphase » mise sur l'adjectif et sur le caractère indéfini du référent, il est possible de répliquer *bir* et d'obtenir des syntagmes polyindéfinis à adjectif antéposé : *bir güzel bir kır*, une belle fille (structure IND A IND N).

1.5 Scénario historique

On a vu que la structure est déjà présente au XIXème siècle : il doit donc s'agir d'un emprunt très ancien. Remonte-t-il à l'époque reculée de l'arrivée des Roms en Europe (auquel cas il aurait concerné l'ancêtre de toutes les variétés actuelles du romani : **hypothèse I**) ? Il serait dans ce cas une rétention, un archaïsme. Ou remonte-t-il seulement à l'époque médiévale (auquel cas il ne concernerait que les Balkans : **hypothèse II**) ? Il serait dans ce cas un (ancien) emprunt.

Suivant Matras (2002 : section 3.2), nous nommons **proto-romani** la variété de romani parlée par les Roms à l'époque de leur vie en Inde et de leur exode, avant leur arrivée au Moyen-Orient et en Europe orientale. Il s'agit d'une langue indienne de l'Inde, aux traits suffisamment distincts des autres idiomes indo-aryens pour être classifié comme une entité propre (Matras 2002 : 18). En l'absence d'attestation écrite, il est toutefois impossible de la localiser précisément dans le temps et dans l'espace : cette langue est une hypothèse de travail. On la reconstruit à partir des autres langues indo-aryennes attestées, anciennes ou modernes, ainsi que des variétés actuelles du romani. On peut également dresser une comparaison avec d'autres langues indo-aryennes diasporiques situées entre l'Inde et l'Europe (Matras 2002 : 19), comme le domari (Matras 1999b, Matras 2012, Herin 2012) et le lomavren (Boerger 1984).

Nous nommons **vieux romani** (*Early Romani*) la variété de romani parlée par les Roms entre l'époque de leur arrivée en Europe et celle de leur dispersion à travers l'Europe, qui a provoqué l'éclatement de la langue en nombreuses variétés géographiques. On situe cette période durant l'empire byzantin, entre le XIème siècle (ou IXème-XIème siècle selon Matras 2002 : 19) et le XIIIème-XIVème siècle). Le passage ou l'arrivée des Roms est en effet attestée ensuite dans divers pays

d'Europe : en Serbie en 1348, en Bohème en 1399, en Basse-Saxe en 1407, en Hesse 1414, à Meissen en 1416, puis en 1417 à Zurich, Magdebourg, Lübeck, Hambourg et en Poméranie en 1417 (Hohmann 1990), en 1419 à Châtillon en Dombes (aujourd'hui Châtillon sur Chalaronne) et à Mâcon, en 1427 à Paris (Courthiade 2013 : 568). Le vieux romani se caractérise par un grand impact morphologique, syntaxique et lexical du grec byzantin, dont ont hérité toutes les variétés contemporaines (Matras 2002 : section 3). L'influence majeure qui nous intéresse est l'apparition, en romani, d'un article défini (cf. section 1.3.1.1 du chapitre I, p. 107).

Le Tableau 104 représente le scénario d'emprunt de la polydéfinitude en romani selon l'hypothèse I : la polydéfinitude a été conservée, depuis le stade du vieux-romani, dans les Balkans et l'Europe du centre-est. C'est la perte de la polydéfinitude qui serait une innovation en Europe du nord-est (de même que la perte de l'article défini qui a suivi) et dans toute l'Europe de l'ouest.

Tableau 104 : Scénario d'une rétention (hypothèse I)

Variété diachronique		Article défini	Polydéfinitude
Proto-romani		×	×
Vieux-romani		✓	✓
Romani moderne :	Balkans, Europe du centre-est	✓	✓
	Europe du nord-est	×	×
	Europe de l'ouest (nord, centre et sud)	✓	×

Le Tableau 105 représente le scénario selon l'hypothèse II : la polydéfinitude est acquise au Moyen-Âge dans les Balkans et l'Europe du centre-est. Il s'agirait d'un emprunt propre à ces régions, datant d'une époque antérieure au XIXème siècle.

Tableau 105 : Scénario d'un ancien emprunt (hypothèse II)

Variété diachronique		Article défini	Polydéfinitude
Proto-romani		×	×
Vieux-romani		✓	×
Romani moderne :	Balkans, Europe du centre-est	✓	✓
	Europe du nord-est	×	×
	Europe de l'ouest (nord, centre et sud)	✓	×

En l'absence d'attestation écrite, il nous est pour le moment impossible de trancher avec certitude entre ces deux scénarii. Pour aller plus loin, il nous faut prendre en compte une autre particularité de la structure polydéfinie, qui est absente des autres langues indo-aryennes anciennes ou modernes : la postposition de l'adjectif.

2. Origine de la postposition de l'adjectif

Dans la mesure où la réplication de l'article défini comporte la postposition de l'adjectif, on pourrait être enclin à penser que la postposition de l'adjectif est historiquement première par rapport à la réplication de l'article. Mais ce n'est pas nécessairement le cas, dans la mesure où, dans notre corpus, les deux phénomènes ne sont pas pratiqués par les mêmes classes d'âge et par les mêmes variétés. De plus, une langue comme le grec permet DEF N DEF A sans permettre DEF N A: les deux ne vont donc pas toujours de pair.

2.1 Problématique

On sait que la place canonique de l'adjectif est avant le nom (cf. section 1.4.3.1, p. 138) dans toutes les langues indo-aryennes modernes (Masica 1991 : 370) dont le romani (Matras 2002 : 165-167). Or on sait maintenant que l'adjectif peut être postposé au nom dans toutes les variétés d'Albanie, en syntagme monodéfini comme polydéfini. La postposition de l'adjectif est une quasi condition de la polydéfinitude en romani d'Albanie. Depuis **quand** est-ce le cas ?

I'. Un premier scénario serait qu'elle ait été acquise par le vieux romani (*Early Romani*) puis héritée en romani albanais contemporain. Toutes les autres variétés du romani en auraient hérité, certaines l'ayant par la suite perdue.

II'. Un deuxième scénario serait qu'elle ait été acquise par le seul romani balkanique durant l'époque médiévale. Les autres variétés d'Europe ne l'auraient donc pas connue.

III'. Un troisième scénario serait qu'elle ait été acquise récemment dans la seule Albanie. Les autres variétés du romani ne l'emploieraient donc pas.

Est-il possible de comprendre **comment** le romani d'Albanie a acquis la postposition de l'adjectif ?

A'. Une première hypothèse serait qu'il s'agisse d'une innovation interne, qui a une fonction particulière et remplit un besoin linguistique.

B'. Une deuxième hypothèse serait qu'il s'agisse d'un emprunt à une autre langue. Reste à savoir laquelle ou lesquelles, comment et quand s'est effectué l'emprunt.

Pour répondre à ces questions, il est nécessaire de combiner l'histoire de la langue et sa géographie. Nous appliquerons les critères méthodologiques proposés par Adamou (2016 : 61) pour déterminer si le phénomène est dû au contact de langues :

1. Décrire le fait linguistique dans les vernaculaires. (chapitres IV et V)
2. Rechercher le fait linguistique en diachronie. (section 2.2, p. 447)

3. Repérer le fait linguistique dans les autres langues de la même famille. (sections 2.3 p. 447, 2.4.1.1 p. 452, et 2.4.2.1 p. 455)
4. Evaluer la rareté typologique du fait linguistique. (sections 2.4.1.2 p. 455, et 2.4.2.2 p. 455)
5. Identifier la (les) langue(s) source(s) et attester le fait linguistique. (section 2.5, p. 457)
6. Tracer les contacts effectifs entre les locuteurs, région par région. (chapitres IV et V)
7. Identifier le processus. (sections 2.5, p. 457, et 3 p. 462)

2.2 Histoire : le romani albanais ancien

Le Tableau 95 montre que le romani albanais du début du XXème siècle connaissait déjà la postposition de l'adjectif dans les syntagmes définis. Elle était d'ailleurs moins fréquente en syntagme monodéfini que polydéfini.

2.3 Géographie : le romani des autres régions

2.3.1. Descriptions d'autres variétés

Nous avons vu en section 1.4.3.2 du chapitre I, p. 139, que la postposition de l'adjectif a été relevée par les travaux suivants :

- Deman (2005 : 77) : la postposition de l'adjectif est non-marquée en romani ***vlax* de Bogota**, dans tous les syntagmes (défini, défini complexe, polydéfini), l'antéposition étant possible en cas d'« emphase ».
- Soravia (1972 : 38) : la postposition de l'adjectif est non-marquée en romani des **Abruzzes**.
- Leggio (2011 : 103) : la postposition de l'adjectif est le choix « par défaut » en romani de **Molise**, tel que recensé par le corpus RMS (codes IT-007 et IT-010).
- Boretzky (1993 : 171) : la postposition existe en **bugurdži du Kosovo**, en syntagme indéfini, défini possessif ou polydéfini.
- Boretzky (1994) : la postposition existe en **kalderaš du Banat serbe**, en syntagme indéfini, défini possessif, monodéfini ou polydéfini.
- Igla (1996 : 166) : la postposition est présente en romani d'**Ajia Varvara**, en syntagme indéfini et polydéfini.
- Matras (2004 : 97) : la postposition est présente en **romacilikanes de Parakalamos**, en syntagme indéfini et polydéfini.
- Leggio (2011 : 102-104) : la postposition de l'adjectif existe en romani de **Mitrovica** au Kosovo, tel que parlé par des locuteurs immigrés en Sicile. D'après les extraits analysés, on la trouve en syntagme introduit par article défini, article zéro, quantifieur, et possessif.
- Adamou (2016 : 161) : la postposition est présente mais très minoritaire et marquée, en romani de **Thrace**.

La présence de la postposition dans de si nombreuses variétés, même en syntagme défini invalide notre **hypothèse III'** selon laquelle elle ne serait apparue que dans la seule Albanie.

2.3.2. Données roumaines contemporaines

Nous avons trouvé, dans les quatre premiers récits de vie de Cioabă (2006a), les structures suivantes avec adjectif postposé :

- **DEF N A :** dix occurrences (contre sept occurrences de DEF A N) → *o stàto ʒermàno*, l'État allemand, *anθ-e dukh kardiakë*, dans la douleur cardiaque, *o punro phago*, le pied cassé, *i munca agricolă*, le travail agricole, *o Piculo baro*, le grand Piculo, *o śil baro*, le grand froid, *i neàmcuri laćhe*, les Allemands gentils, *o fòro baro*, la grande ville…
- **Ø N A :** vingt-quatre occurrences (contre quatre occurrences de Ø A N) → *kriminàlo baro*, un grand criminel, *plànuri barë*, de grands plans, *ćhej bari*, une grande fille, *kolompìrea bilonde*, des pommes de terre non-salées, *gav bulgërìsko*, un village bulgare, *paj ćikalo, melalo*, de l'eau boueuse, sale, *jag bari*, un grand feu…
- **IND N A :** vingt-et-un occurrences (contre cinq occurrences de IND A N) → *ëkh kotor baro*, un grand morceau, *ëkh solùcia ćaći*, une vraie solution, *ëkh romni phari*, une femme enceinte, *ëkh muzèo romano*, un musée rom, *nìśte śax śuklo*, des choux acides, *ëkh gurumni muli*, une vache morte, *ëkh śwaoo tërno*, un jeune garçon, *anθ-ëkh gaw òthe većìno*, dans un village voisin là…
- **N A POSS** : une occurrence (pas d'occurrence de POSS N A ni de POSS A N) → *Roma vrèdnići tumarë*, vos Roms dignes.
- **DEF N DEM A :** une occurrence (pas d'occurrence de DEM N A ni de DEM A N) → *e braval kodea śudri,* ce vent froid.
- **NUM/QUANT N A :** huit occurrences (pas d'occurrence de NUM/QUANT A N) → *ëkh śel gàlbi barë*, cent grosses pièces d'or, *biś thaj panʒ gàlbi barë*, vingt-cinq grosses pièces d'or, *duj gàlbi barë*, deux grosses pièces d'or, *trin ćheja barea*, trois grandes filles…

Non seulement la postposition de l'adjectif existe dans ces variétés de romani de Roumnie, mais elle semble plus fréquente que l'antéposition, quel que soit le déterminant qui introduit le syntagme.

2.3.3. Données roumaines anciennes

Le Tableau 106 présente quelques occurrences, la traduction donnée par l'autrice en roumain et en anglais, ainsi que notre propre traduction en français. Il montre que la postposition de l'adjectif est répandue de très longue date dans toute la zone roumanophone. Gardons en mémoire que ces occurrences sont issues de chants et de contes, donc d'une parole élaborée et non-spontanée. Comme en roumain contemporain, on observe que l'adjectif peut se postposer quel que soit le déterminant introducteur : article défini, possessif, démonstratif, article indéfini, article zéro, quantifieur.

Tableau 106 : Exemples de syntagmes avec adjectif antéposé dans les textes collectés par Barbu Constantinescu

Dialecte	Comté	Texte	Structure	romani	roumain	anglais	français
ursari *ricinearísko ursăresc*	Bacău	155	DEF N A	*o biau romano*	*nuntă țigănească*	*Gypsy wedding*	le mariage rom
	Prahova	48	DEF N A	*e blaval șudri*	*vântul rece*	*the cold wind*	le vent froid
	Roman	67	Ø N (SGEN) A	*patrin șahăski parno*	*foaie de varză albă*	*white cabbage leaf*	feuille de chou blanc
		68	DEF N A	*o ć<h>eaoró melahnó*	*băiatul oacheș*	*the swarthy cute chap*	le jeune garcon sombre (sale)
	Dâmbovița	71	Ø N A	*vatáfo baró*	*vătaf mare*	*great bailiff*	un grand *vataf*
		125	DEF N A	*o brîșind șudroró*	*ploaie recișoară*	*a coldish rain*	la pluie froide
laiesi *lăiășisko lăieșesc*	Dambovița	6	DEF N A	*o mărǵano rupuno*	*mărgeanul de argint*	*the silvery coral*	le corail en argent
		7	IND N A	*ek kor parni*	*un gât alb*	*a fair neck*	un cou blanc
		49	Ø N A	*patrin zeleno*	*frunză verde*	*green leaf*	feuille verte
	Ilfov	52	DEM N A	*koi șei bari*	*acea fată mare*	*that maiden*	cette grande fille
		85	DEF N A	*ăl fesoya loloró*	*cu fesul roșior*	*with pretty red fez*	avec le joli fez rouge
		93	PREP DEF N A	*ko-u plai baro*	*la dealul mare*	*to the high hill*	à la grande colline
			DEM N A	*kola moleate nevi*	*la acea cârciumă nouă*	*at that new pub*	dans cette nouvelle taverne
	București	89	Ø N A A	*k< ’>ăr baro somnakuó*	*casă mare de aur*	*golden big house*	grande maison en or
	Romanați	91	Ø N A	*mol guli*	*vin dulce*	*sweet wine*	vin doux
			PREP DEM N A	*ande koi stikla bari*	*în acea sticlă mare*	*in that big bottle*	dans cette grande bouteille
	Teleorman	165	NUM N A	*trin șeya barea*	*trei fete mari*	*three big girls*	trois grandes filles
kalderaš	București	55	IND N A	*ek rai baro*	*un boier mare*	*a big boyar*	un grand seigneur

Dialecte	Comté	Texte	Structure	romani	roumain	anglais	français
kikavearisko căldărăresc	Iași	66	PREP DEF N A	*pala-i kurva marime*	*după curva cea spurcată*	*to the soiled whore*	après la catin souillée
tismani *tismănarisko tismănăresc*	Gorj	83	DEF N A	*le fesosa loloró*	*cu fesul roșu*	*with pretty red fez*	avec le joli fez rouge
		276	Ø N A	*abeáu romanó*	*nuntă țigănească*	*Gypsy wedding*	mariage rom
			DEF N A	*lă gadesa melaló*	*cu cămașa murdară*	*with the dirty shirt*	avec la chemise sale
		117	PREP DEF N A	*and-ă vășănda tarnă*	*la pădurile tinere*	*to the young forests*	dans les jeunes forêts
		169	DEF N A	*i bori șukar*	*nora frumoasă*	*the daughter in law, beautiful*	la jolie belle-fille
	Craiova	32	Ø N A	*șei bari*	*o fată mare*	*a maiden*	grande fille
			DEM N A	*kou bal șukar*	*acel păr frumos*	*that beautiful hair*	ce beau cheveu
			POSS N A	*muro bal somnakuno*	*al meu păr de aur*	*my golden hair*	mon cheveu doré
zlatari *zlătarisko zlătăresc*	Craiova	199	PREP DEM N A	*pe koi poyana barí*	*pe acea poiană mare*	*on that broad glade*	dans cette vaste clairière
			DEF N A	*ăl belea somnakuné*	*furcile de aur*	*with golden beams*	les fourches en or
			Ø N A	*kilodoreá rupuné*	*târsânele de argint*	*with silver transversal beams*	ces poils de chèvre argentés
	Turnu Severin	250	Ø N A	*rom biboldó*	*țigan nebotezat*	*an unbaptized Gypsy*	Rom non baptisé
			QUANT N A	*sa mol guli*	*tot vin dulce*	*sweet wine*	tout le vin doux
	București	86	Ø N A	*rom lać<h>ó*	*om bun*	*a good man*	un brave homme
allemand *nemțisko nemțesc*	?	2	DEF N A	*ăl koverći rupune*	*covercile de argint*	*with silver beams*	faisceaux en argent
			DEF N A	*ăl beranda somnatune*	*coardele de aur*	*with golden strings*	cordes en or

2.3.4. Autres données

2.3.4.1. Macédoine et Kosovo

Les témoins de Macédoine et du Kosovo serbophone que nous avons pu enregistrer pour notre corpus semi-spontané ont produit les ordres suivants (voir Tableau 66, p. 309) :

- Macédoine : DEF A N à 94,6% et DEF N A à 5,4 %
- Kosovo : DEF A N à 88,89 % et DEF N A à 11,11 %

Le Tableau 69, p. 326, montre que dans les syntagmes déterminés par un autre déterminant (démonstratif, possessif, article indéfini, quantifieur, numéral, article zéro), la tendance est la même que dans les syntagmes définis : l'adjectif y est toujours antéposé, de manière presque catégorique (86,67 %). Ainsi la postposition de l'adjectif est-elle possible en syntagme monodéfini en Macédoine et au Kosovo, mais très rare, et ce même dans le contexte contrastif de la tâche.

2.3.4.2. Grèce

Komotini

Le conte *The louse and the Rom*, la puce et le Rom (Adamou 2006a), présente une occurrence de syntagme indéfini avec adjectif postposé :

(273) Romani (Adamou 2006a : ligne 98)

dikhel	*ek*	*tanava*	*romni*	*phuri.*
see.3SG	one	CLASS	rom.F	old

'He sees an old Rom woman.'

Le conte *The coward and the giants*, le lâche et les géants (Adamou 2006b), présente deux occurrences de syntagme indéfini avec adjectif postposé, l'un introduit par l'article indéfini, l'autre par l'absence d'article.

(274) Romani (Adamou 2006b : ligne 51)

jama	*lengo*	*jek*	*payuri*	*gene*	*jek*	*bardako*		
but	3PL.GEN.M	one	gourd	again	one	glass		
kaj	*te*	*les*	*le*	*sar*	*si*	*jek*	*kher*	*baro.*
where	COMP	take.3SG	3SG.OBL.M	how	be.3SG	one	house	big

'But, their gourd was like a big house.'

(275) Romani (Adamou 2006b : ligne 71)

gadal	*dikhel*	*ambrola*	*ʃukar.*
this.way	see.3SG	pears	nice

'He sees the nice pears.'

Pour compléter avec précision ce tour d'Europe des variétés de romani, une étude approfondie du corpus RMS serait nécessaire. Il faudrait relever l'ensemble des phrases du questionnaire comprenant un syntagme avec adjectif épithète, puis observer leur traduction dans tous les dialectes du monde recensés. On distinguerait les dialectes qui postposent ou antéposent l'adjectif, et surtout, ceux qui pratiquent les deux et dans quels cas. Il conviendrait également de comparer les phrases entre elles, afin de comprendre quels facteurs déterminent la place de l'adjectif, dans quels dialectes. On prendrait en considération la nature du déterminant (défini ou indéfini) qui introduit les syntagmes concernés. Ceci pourra l'objet d'une recherche ultérieure.

2.4. Scénario géographique

2.4.1. Dans les Balkans

2.4.1.1. Le romani

La postposition de l'adjectif a été observée, à notre connaissance, en Albanie, en Grèce, au Kosovo, en Serbie et en Roumanie (cf. section 2.3, p. 447). Les publications attestent également de sa présence en Italie et en Colombie. Une telle extension rend difficile l'**hypothèse A'** de l'innovation interne. On préférera pour l'instant opter pour l'**hypothèse B'** de l'emprunt à une langue de contact. Un grand nombre d'études montre d'ailleurs que l'ordre des mots est parmi les phénomènes linguistiques les plus sensibles au contact de langues (Heine 2008 : 34). Le vieux romani a par exemple changé l'ordre des mots à l'échelle de la phrase, passant du SOV hérité des langues indo-aryennes à l'ordre SVO (éventuellement VSO) du grec et des langues balkaniques (Matras 1996 : 64).

2.4.1.2. Les autres langues

Quelle pourrait être la ou les langues d'emprunt ? Le Tableau 107 montre quelles langues postposent ou antéposent l'adjectif de façon canonique.

Tableau 107 : Présence de la postposition de l'adjectif dans les autres langues des Balkans

	Article défini	Polydéfinitude	**Postposition A**
Albanais (tosque & guègue)	✓	(✓)	✓
Roumain, moldave	✓	✓	✓
Aroumain, mégléno-roumain	✓	✓	✓
Grec	✓	✓	×
Bulgare	✓	×	×
Macédonien, serbe du sud-est	✓	×	×

	Article défini	Polydéfinitude	**Postposition A**
Serbo-croate-bosniaque-monténégrin	×	×	×
Slovène	×	(✓)	×
Turc rouméliote	×	×	×
Hongrois	✓	×	×

Toutes les langues des Balkans antéposent l'adjectif sauf l'albanais et les langues romanes. On peut en déduire que les variétés albanaises et roumaines de romani, qui postposent le plus volontiers l'adjectif, ont été influencées par l'albanais ou les langues roumaines. Elles ont pu par la suite influencer les autres variétés de romani, et au Kosovo (où l'on parle albanais) et au Banat serbe (où l'on parle valaque). Cela ne signifie pas que les autres langues des Balkans ne puissent postposer l'adjectif dans un certain but communicatif et dans certaines tournures. Nous savons par exemple que c'est le cas des syntagmes grecs indéfinis (cf. section 2.2.2.4, p. 158) et polydéfinis. Il est difficile de savoir si l'emprunt est dû à une langue unique ou non : il est possible que les variétés de romani aient chacune effectué un emprunt à leur langue de contact locale.

Albanais

On peut consulter Androutsopoulou (2001), Tomić (2006 : 191-194), Campos (2009) et Feuillet (2012). L'adjectif est canoniquement postposé en albanais (Androutsopoulou 2001 : 163-164). Il peut être antéposé lorsque l'on veut le « focaliser » (Campos 2009 : 1019), lui conférer une « interprétation de focus » (Androutsopoulou 2001 : 164), de l'« emphase » ou du « contraste », notamment au superlatif (Lyons 1999 : 75). Même antéposé, un adjectif de classe 2 (ou un numéral en fonction adjectivale) est toujours accompagné de son article. De plus, en tant que premier lexème du syntagme, c'est alors lui qui porte l'article défini clitique postposé : le nom reste nu.

Roumain

On peut consulter Campos (2005), Feuillet (2012) ainsi que Cinque (2010 : 102-104), Plank (2003b : 346-349) et Tomić (2006 : 128-134). L'adjectif est canoniquement postposé en roumain (Plank 2003b : 346, 347, Tomić 2006 : 127-128, Lyons 1999 : 75). Certains adjectifs sont exclus de l'antéposition, comme *bun*, bon, selon Tomić (2006 : 128) … mais l'on trouve des exemples contraires dans les études, comme *bunii prieteni*, les bons amis (Cinque 2010 : 102). La plupart des adjectifs postposés peuvent entrer dans la structure polydéfinie. Il existe quelques exceptions, telles que les adjectifs relationnels (*american*, américain), postposés mais qui ne peuvent entrer dans une structure polydéfinie (Cornilescu et Nicolae 2012 : 1092).

Quelques adjectifs sont canoniquement antéposés, tels que les adjectifs « intensionnels » (*biet*, pitoyable) : ils ne peuvent alors entrer dans la structure polydéfinie (Cornilescu et Nicolae 2012 : 1092). Selon (Alexiadou 2014 : 54-55), l'antéposition de l'adjectif suit les caractéristiques sémantiques des autres langues romanes, telles qu'exemplifiées par l'italien dans (Cinque 2010). L'adjectif antéposé ne peut être interprété que de manière non-restrictive, tandis que l'adjectif postposé peut être interprété de manière soit restrictive soit non-restrictive.

En syntagme polydéfini, l'adjectif est postposé avec l'article *cel* (*soldatul* *cel bun*). L'adjectif postposé ne peut être interprété que de manière restrictive (Figure 78). L'antéposition de la « sous-partie adjectivale » (*cel bun* *soldatul*, le bon soldat) est marquée et s'interprète comme une « emphase » sur l'adjectif ou une « différenciation sémantique ».

Exemple	**Structure**	**Interprétation**
legi=le importante	N=DEF A	non-restrictive
importante=le legi	A=DEF N	restrictive ou non-restrictive
legi=le cele importante	N=DEF ART A	restrictive

Figure 78 : Interprétation et ordre des constituants du SN « les lois importantes » en roumain

Aroumain

On peut consulter Campos (2005) et Tomić (2006 : 169-171) pour l'**aroumain** de Thessalie (Grèce). L'adjectif est par défaut antéposé dans cette langue. La postposition est possible de manière optionnelle dans les syntagmes indéfinis à article zéro (Tomić 2006 : 169). Elle est possible de manière obligatoire dans les syntagme polydéfinis.

Mégléno-roumain

Comme en aroumain, l'adjectif est par défaut antéposé (Tomić 2006 : 155-156). Il peut cependant être postposé en cas d'emphase (Tomić 2006 : 155), ce qui n'est pas le cas de l'aroumain.

Langues slaves

Les langues slaves antéposent canoniquement l'adjectif. Pour le **bulgare**, on peut consulter Feuillet (2012) et Tomić (2006 : 91-94), pour le **macédonien** Foulon (1997) et Tomić (2006 : 118-125). Pour le **serbo-croate**, on peut consulter Aljović (2010), Thomas et Osipov (2012), Feuillet (2012) et Marušič et Žaucer (2008), ainsi que les notes de Cinque (2010 : 99-102) et Tomić 2006 : 110-111). Pour le **slovène familier**, on peut consulter Marušič et Žaucer (2008).

Turc rouméliote

On peut consulter Friedman (2003a) et Friedman (2006).

Hongrois

On peut consulter Moravcsik (2003) et Szendrői (2010).

2.4.2. Hors des Balkans

2.4.2.1. Le romani

Sans étude large portant sur l'ensemble des variétés, par exemple via le corpus RMS, il nous est pour le moment impossible de cartographier la postposition de l'adjectif en Europe. Le Tableau 108 propose des hypothèses, fondées notamment sur la présence ou non de langues romanes de contact.

Tableau 108 : Hypothèse de présence de postposition de l'adjectif en Europe

	Article défini	Polydéfinitude	**Postposition A**
Balkan	✓	✓	✓/×
Vlax	✓	✓	✓
Centre	✓	✓	×
Nord-est	×	×	×
Nord-ouest + Italie + Espagne	✓	×	✓/×

2.4.2.2. Les autres langues

Le Tableau 109 présente quelles langues postposent ou non l'adjectif de façon canonique.

On constate que la plupart des langues antéposent l'adjectif, ce qui n'a aucune raison d'inciter le romani à le postposer ; cela permet de conserver la structure canonique DEF A N. Seules les langues romanes, les langues celtiques, le maltais (sémitique) et le basque (isolat) postposent canoniquement l'adjectif. Ce constat est cohérent avec celui de la postposition de l'adjectif en Italie (Abruzzes) et en Colombie (hispanophone), hors des Balkans. On peut établir une comparaison avec une langue slave à adjectif antéposé (A N et IND A N) qui s'est mise à postposer l'adjectif (N A et IND N A) par contact avec une langue romane : le croate molisien parlé au sud de l'Italie, en contact intense avec l'italien depuis un demi-millénaire (Breu 1996). Il s'agit d'un cas de *pragmatic unmarking* selon Heine (2008 : 44). Pour l'ensemble des langues, on peut consulter Alexiadou (2014) qui étudie la position de l'adjectif et la structure sous-jacente des syntagmes adjectival et nominal, en comparant notamment langues scandinaves, grec, anglais, italien, français, hébreu, albanais, roumain, slovène et chinois.

Tableau 109 : Présence de la postposition de l'adjectif dans les principales langues d'Europe

Zone	Langue	Article défini	Polydéfinitude	Postposition A
Est	Russe	×	×	×
	Ukrainien	×	×	×
	Biélorusse	×	×	×
	Polonais	×	×	×
	Tchèque	×	×	×
	Slovaque	×	×	×
Nord	Letton	×	×	×
	Lituanien	×	×	×
	Estonien	×	×	×
	Finlandais	×	×	×
	Suédois	✓	✓	×
	Norvégien	✓	✓	×
	Féroïen	✓	✓/×	×
	Islandais	✓	×	×
Ouest	Danois	✓	×	×
	Frison	✓	×	×
	Allemand	✓	×	×
	Bavarois	✓	✓	×
	Yiddish	✓	✓	×

Zone	Langue	Article défini	Polydéfinitude	Postposition A
Ouest	Néerlandais	✓	×	×
	Flamand	✓	×	×
	Anglais	✓	×	×
	Gaélique d'Irlande, d'Ecosse	✓	×	✓
	Gallois	✓	×	✓
	Breton	✓	×	✓
	Français	✓	×	✓/×
Sud	Basque	(✓)	×	✓
	Catalan	✓	×	✓/×
	Espagnol	✓	×	✓/×
	Galicien	✓	×	✓/×
	Portugais	✓	×	✓/×
	Occitan	✓	×	✓/×
	Franco-provençal	✓	×	✓/×
	Italien, corse	✓	×	✓/×
	Sarde, sicilien	✓	×	✓/×
	Maltais	✓	✓	✓
	Chypriote	✓	✓	×
	Turc	×	×	×

Pour le **russe**, on peut consulter Babby (2010) et Cinque (2010 : 108-110), pour le le **lituanien** Penanros (2015). Pour les langues **scandinaves**, on peut lire Dahl (2004) et surtout Marit (2005). Alexiadou (2014 : 113) note que la postposition de l'adjectif était possible et relativement productive en **vieil-islandais**.

Pour l'**allemand**, on pourra consulter Plank (2003b : 340-341, 366-367, 372-373), Leu (2009) et Cinque (2010 : 111). Le **bavarois** fait l'objet d'analyses dans Plank (2003b : 374-375), Alexiadou (2014 : 120) ainsi que de remarques dans Lekakou et Szendrői (2012 : 143-144). Pour le **yiddish**, on peut consulter Plank (2003b : 345-346), Taube (2012 : 467-468) et Jacobs (2005 : 174). Le livre de Cinque (2010) fait référence pour l'adjectif en **anglais**, dont le fonctionnement est censé valoir pour l'ensemble des langues germaniques. Pour une étude comparative de ces langues, on peut consulter également Schoorlemmer (2012).

Toutes les **langues celtiques** postposent l'adjectif (Ball et Müller 2010 : 16), ce qui est attesté depuis le vieil irlandais (Ball et Müller 2010 : 60). Il existe toutefois quelques adjectifs antéposés, pour la plupart les mêmes d'une langue à l'autre : gallois *hen*, vieux, breton *gozh*, vieux (Borsley et Roberts 1996 : 38-39).

Le livre de Cinque (2010) fait référence pour l'adjectif en **italien**, dont le fonctionnement est censé valoir pour l'ensemble des langues romanes. Pour le **français**, on peut consulter Forsgren (1978) Wilmet (1981a), Wilmet (1981b), Forsgren (1997), Bouchard (1998), Abeillé et Godard (1999), Laenzlinger (2000) et Laenzlinger (2005), Nølke (2001), Thuilier (2012) et Thuilier (2015).

Le **maltais** postpose l'adjectif, à l'instar des autres langues sémitiques[1]. On peut consulter l'article de Fabri (2001), les notes de Plank (2003b : 347, 350-351) ainsi que Cinque (2010 : 98-99). Certains adjectifs, notamment empruntés, sont toutefois antéposés.

Pour le **turc**, on pourra consulter (Plank 2003b : 351-352) : l'adjectif est strictement anténominal. En cas de focus contrastif, il peut se déplacer avant l'article indéfini : *güzel bir kır*, une belle fille (structure A IND N).

2.5. Scénario historique

On a vu que la structure existait déjà au XIXème siècle : il doit s'agir d'un emprunt ancien. À quand remonte-t-il ? Elle ne peut remonter à l'époque de l'arrivée des Roms en Europe, celle du proto-romani ni du vieux romani car c'est le grec, langue à antéposition, qui était la langue de contact principale dans l'empire byzantin. Cela invalide l'**hypothèse I'** d'une rétention, d'un archaïsme. On ne peut donc plus opter que pour l'**hypothèse II'**, celle d'un ancien emprunt à l'époque médiévale. Il

[1] L'adjectif est postposé au nom en hébreu (Alexiadou 2014 : 79-90) et en arabe (Plank 2003b : 355-356). A l'exception des quantifieurs et de l'article défini, tous les modifieurs sont postposés au nom : démonstratifs, numéraux ordinaux, adjectifs, syntagmes génitivaux, syntagmes prépositionnels (Yatziv-Malibert 2015).

nous faut toutefois la modifier : rien n'indique que l'emprunt ne concerne que les Balkans. Il peut concerner, par exemple, la péninsule ibérique ou l'Italie.

Le Tableau 110 représente le scénario d'emprunt selon l'hypothèse II' : la postposition de l'adjectif est acquise au Moyen-Âge dans les Balkans et en Europe de l'ouest. Il s'agirait d'un phénomène propre à ces régions, mais datant d'une époque reculée, avant le XIXème siècle.

Tableau 110 : Scénario d'un ancien emprunt (hypothèse II')

Variété diachronique		Article défini	Polydéfinitude	**Postposition A**
Proto-romani		×	×	**×**
Vieux-romani		✓	×	**×**
Romani moderne :	Balkans, Europe du centre-est	✓	✓/×	**✓/×**
	Europe du nord-est	×	×	**×**
	Europe de l'ouest (nord, centre et sud)	✓	✓/×	**✓/×**

2.5.1. Problème de chronologie des acquisitions

Une question se pose pour nous : l'emprunt de la postposition s'est-il produit après ou en même temps que celui de la polydéfinitude ?

- Dans le scénario de l'hypothèse I et II', les variétés auraient acquis **d'abord la polydéfinitude** (*toutes* les variétés de romani, qui ne l'ont pas forcément conservée par la suite), **puis la postposition de l'adjectif** (*seules* certaines variétés), en syntagme mono- et polydéfini.
- Dans le scénario de l'hypothèse II (et II'), la variétés auraient acquis **à la même époque** la polydéfinitude (les variétés des Balkans et d'Europe du centre-est, via le grec) et la postposition de l'adjectif (seules certaines variétés des Balkans et de l'Europe du centre-est, via plusieurs langues de contact possibles).

Il convient toutefois de noter que cela dépend des langues de contact, donc des régions :

1) l'emprunt de la postposition de l'adjectif ne suppose pas l'acquisition préalable de la polydéfinitude en **Europe de l'ouest**, où les langues romanes, celtiques et le basque postposent l'adjectif en tous syntagmes et où la polydéfinitude n'existe pas.

Si la polydéfinitude a été acquise *avant* la postposition de l'adjectif (scénario I et II'), alors cette dernière a pu en être facilitée, car cela signifie qu'une forme de postposition de l'adjectif existait déjà en langue romani, même si ça n'était qu'en syntagme polydéfini.

Le scénario II et II' de l'emprunt *concomitant* est quant à lui impossible dans cette région : les populations roms concernées ne pouvaient pas se trouver au même moment dans les Balkans et en Europe de l'ouest.

2) l'emprunt de la postposition de l'adjectif ne suppose pas l'acquisition préalable de la polydéfinitude en **zone roumanophone**, car les langues roumaines postposent l'adjectif aussi bien en syntagme monodéfini que polydéfini ou indéfini.

Dans cette zone, on ne peut juger si la polydéfinitude a été acquise (via le grec) *avant* la postposition (via les langues roumaines) et a alors facilité cette dernière (scénario I et II').

… ou bien si les deux structures ont été acquises *en même temps* en se favorisant mutuellement (scénario II et II').

3) l'emprunt de la postposition de l'adjectif suppose l'acquisition préalable de la polydéfinitude dans les zones **hellénophone, albanophone et maltaise**, où le grec, l'albanais et le maltais postposent l'adjectif seulement en syntagme « polydéfini ».

Seul le scénario I et II' est possible pour le romani de ces régions : *d'abord* la polydéfinitude, ensuite la postposition de l'adjectif, facilitée par l'existence de DEF N DEF A, polydéfinitude à adjectif postposé.

Pour supposer que le scénario II et II' soit possible, il faudrait supposer que la postposition de l'adjectif ait fait l'objet non d'un emprunt, mais d'une *innovation interne*. Les variétés concernées auraient eu l'idée de postposer directement l'adjectif, sans l'articuler, alors même que cela n'existe ni en grec ni en albanais.

2.5.2. Solution de chronologie

On comprend que le double scénario I et II' est plus plausible dans presque tous les cas. Il faut postuler une séparation entre l'acquisition de la polydéfinitude (à adjectif antéposé ou postposé) et la postposition de l'adjectif (en syntagme monodéfini, polydéfini ou indéfini). Nous proposons la solution présentée dans la Figure 79 ci-dessous.

- Le romani a acquis l'article défini et la polydéfinitude par contact avec le grec à l'époque byzantine. Les groupes se sont ensuite dispersés et les variétés géographiques se sont créées.
- Certains groupes sont arrivés dans des zones où les langues de contact ne présentent pas d'article défini : ils ont commencé par perdre la polydéfinitude avant de perdre progressivement l'article défini lui-même. Il se trouve également que ces langues de contact antéposent l'adjectif : ces groupes n'ont donc pas acquis la postposition de l'adjectif.

- D'autres groupes sont arrivés dans des zones où les langues de contact présentent un article défini mais pas de polydéfinitude : ils ont perdu l'usage de la polydéfinitude mais ont conservé l'article défini. Selon la place de l'adjectif dans ces langues,
    - certains groupes n'ont pas acquis la postposition de l'adjectif (Allemagne),
    - d'autres groupes ont acquis la postposition de l'adjectif (Italie). Cette acquisition a été d'autant plus favorisée si elle s'est produite à une époque où la polydéfinitude n'était pas encore tout à fait tombée en désuétude.
- D'autres groupes sont restés dans les Balkans ou sont montés vers le centre de l'Europe orientale, où de nombreuses langues de contact présentent aussi bien un article défini qu'une forme de polydéfinitude : ils ont conservé la polydéfinitude aussi bien que l'article défini. On sait qu'ensuite, aux XIXème, XXème et XXIème siècles, de nombreux groupes *vlax* de la zone roumanophone ont migré vers des zones à article défini mais sans polydéfinitude, comme la France, l'Allemagne, la Colombie... Finiront-ils par perdre la polydéfinitude, lorsque le roumain sera devenue leur L2 ancienne, *older L2* (cf. section 2.2.1 du chapitre I, p. 34) Matras 2002 : 195) ? En fonction de la place de l'adjectif dans ces langues,
    - certains groupes n'ont pas acquis la postposition de l'adjectif en syntagme monodéfini (Serbie, Macédoine, Bulgarie). Ils ne postposent donc l'adjectif, aujourd'hui, qu'en syntagme polydéfini DEF N DEF A (structure héritée du vieux romani). On peut supposer que la polydéfinitude elle-même est en perte de vitesse, car il n'en existe aucun équivalent dans la langue de contact, qui ne présente d'article défini (serbo-croate) ni ne postpose l'adjectif (serbo-croate, macédonien, bulgare).
    - d'autres groupes ont acquis la postposition de l'adjectif (Roumanie).
    - d'autres groupes ont acquis la postposition de l'adjectif (Grèce, Albanie) mais l'emploient rarement et dans des structures marquées, car la langue de contact ne présente pas de syntagme monodéfini à adjectif postposé DEF A N, mais le syntagme polydéfini à adjectif postposé DEF N DEF A.

Figure 79 : Scénario historique et géographique pour l'acquisition de la polydéfintiude et de la postposition de l'adjectif dans les différentes variétés de romani

Bien sûr, rares sont les régions d'Europe monolingues, notamment avant le XIXème siècle. Le tableau d'ensemble que nous venons de dépeindre demande à être nuancé, notamment en fonction de la langue de l'interaction sociale au quotidien hors du cercle des proches (mésolecte) et de la langue de prestige et d'interaction avec les autorités (acrolecte) (cf. section 2.2.2 du chapitre I, p. 35). Il convient également de prendre en compte l'échelle considérée : la langue principale du quartier n'est pas toujours la même que de la ville, ni de la région, ni de l'ensemble du pays. La langue de l'école n'est pas non plus toujours celle des écoles du reste du pays. Dans le banat serbe, la langue locale peut être le valaque et non le serbe ; en Voïvodine, elle peut être le hongrois et non le serbe. Au Kosovo, il convient de distinguer la zone serbophone de la zone albanophone (guègue). Pour la Moldavie, on distingue la zone roumanophone/moldavophone de la zone ou vivent les communautés russophones.

La Thrace est une région particulière de Grèce. Adamou (2010) esquisse par exemple un historique du contact linguistique de la communauté rom qu'elle a étudiée. Arrivés dans la région dans les années 1920 suite au traité de Lausanne, les Roms *vlax* de Komotini forment un groupe fermé mais lié aux autres groupes roms vivant dans les autres quartiers de la périphérie ou dans la ville voisine de Xanthi. Ils parlent romani à la maison, mais grec et turc dans les activités professionnelles et commerciales. Le turc est la langue dominante et de prestige, et c'est également la langue de l'école pour les enfants qui sont scolarisés (Adamou 2010 : 164). À tel point que dans certaines familles aisées, la langue romani cesse d'être transmise aux enfants en faveur du turc (*language shift*).

2.5.3. Objections

On peut faire deux objections au tableau d'ensemble que avons dépeint en section 2.5.2, p. 459.

1) Pourquoi la structure DEF A DEF N est-elle si rare dans les variétés connaissant la polydéfinitude ? Pourquoi DEF A DEF N est-il minoritaire par rapport à DEF N DEF A en romani, alors même que le romani a hérité DEF A N et acquis DEF N A bien plus tard, et alors même que le grec utilise aussi bien DEF A DEF N que DEF N DEF A ?
2) Pourquoi la structure DEF N A est-elle si présente et si forte en Albanie contemporaine, alors même que l'albanais ne possède pas cette structure mais seulement une forme similaire à la polydéfinitude DEF N DEF A ?

Pour répondre à la première objection, on peut dire que la structure DEF A DEF N remonte en principe aussi loin que DEF N DEF A. Elle a probablement connu ensuite une perte de vitesse, jusqu'à être quasiment perdue aujourd'hui. Une première explication serait qu'elle fait aujourd'hui double emploi avec la nouvelle structure monodéfinie DEF N A. En effet, DEF A DEF N a pour but d'opérer une focalisation grâce au deuxième article défini. Or la postposition de l'adjectif poursuit le même but, même en syntagme monodéfini DEF N A – surtout pour les variétés où DEF N A est rare et marqué, comme chez les Arlis âgés albanais. La concurrence entre l'ancien DEF A DEF N et le nouveau DEF N A s'est faite au détriment de DEF A DEF N, peu à peu tombé en désuétude. En Grèce au contraire, DEF N A et DEF A DEF N A continuent à être en concurrence. Une deuxième explication serait l'influence de la langue de contact. Ni l'albanais ni le roumain ne connaissent de structure DEF A DEF N, ce qui n'a pas favorisé le maintien de cette structure. Le grec, quant à lui, connaît DEF A DEF N mais non DEF N A : en romani, la concurrence entre les deux structures tourne en faveur du maintien de DEF A DEF N.

En ce qui concerne la deuxième objection, notre étude a montré que la postposition de l'adjectif n'est pas rare ni marquée en romani d'Albanie. Nous avons pu constater que DEF N A n'est pas (plus) une structure marquée en romani d'Albanie, au vu de son usage important chez l'ensemble de nos locuteurs. Si l'on considère l'usage de DEF N DEF A, où l'adjectif est postposé, on peut même dire que la postposition est devenue majoritaire dans la langue, et la place non marquée de l'adjectif dans le contexte de nos tests (68,33% des occurrences). Il en va de même pour les syntagmes indéfinis, où IND N A représente 64% des occurrences. La réponse à la deuxième objection nécessite une approche socio-linguistique.

3. Scénario socio-linguistique

3.1. Types d'études

On distingue deux types majeurs d'études en socio-linguistique :

- les études **transversales** (*cross-sectional research* ou *trend studies*), qui consistent à étudier une communauté qui a déjà fait l'objet d'une précédente recherche, tout en s'efforçant d'en copier la méthode (pour échantillonner, collecter et analyser les données),

- les études **longitudinales** (*longitudinal* ou *panel studies*), qui consistent à suivre et à étudier les mêmes individus ou groupes pendant une période donnée afin de suivre leur évolution (Sankoff 2005 : 1003). Lorsque l'on détecte un changement chez un individu, deux interprétations sont possibles :

    - l'individu réalise un changement qui va à l'encontre du reste de la communauté
    - l'individu réalise un changement qui va dans la même tendance que le reste de la communauté, en train de modifier sa façon de parler (Sankoff 2005 : 1003).

Labov (1994 : 85) a remarqué que les traits morphologiques, lexicaux et discursifs ont plus de probabilité de changer au cours de la vie d'un individu que ceux phonologiques. C'est expliqué également par Meyerhoff (2011 : 197) :

> « *One of the clearest findings emerging from panel studies is that not all linguistic variables behave the same across a speaker's lifespan. As a general rule, a speaker's phonology is more stable than their vocabulary. You will certainly have acquired or learnt new words for concepts and things after the critical period, and we keep acquiring vocabulary throughout our entire lives. In between phonology and lexicon, though, variables behave a bit differently. Some syntactic and morphological variables seem to be treated by speakers as if they were essentially lexical, and so we see the ability for individuals to restructure their systems radically over real time. Some syntactic and morphological variables seem relatively stable though, and this suggests that speakers understand them as being more like phonological variables*. » (Meyerhoff 2011 : 197)

3.2. Types de résultats

Que peut-on déduire d'une variation en fonction de l'âge dans une langue ? Comment articuler changement linguistique au sens historique, pour la communauté, changement linguistique au niveau individuel, pour un locuteur particulier ? Labov (1994) montre qu'il existe quatre configurations possibles (Figure 80).

Table 1 Patterns of change in the individual and the community[a]

Synchronic Pattern	*Interpretation*	*Individual*	*Community*
Flat	1 Stability	Stable	Stable
Regular increase/decrease with age	2 Age grading	Unstable	Stable
Regular increase/decrease with age	3 Generational change (apparent time)	Stable	Unstable
Flat	4 Communal change	Unstable	Unstable

[a]Adapted from Labov, 1994.

Figure 80 : Les quatre configurations de changement linguistique à échelle individuelle et collective (Sankoff 2006 : 111, d'après Labov 1994)

Sankoff et Blondeau (2007 : 562-563) ont approfondi le modèle de Labov en distinguant deux sous-types dans la configuration 2 *age grading* : 2a *age grading* et 2b *lifespan change*. Il en résulte cinq configurations différentes, représentées en Figure 81.

1) la **stabilité** (*stability*) : aucun changement n'est en cours.

2) la **gradation par l'âge** (*age grading*) : des individus isolés changent leur façon de parler et se distinguent par là du reste de la communauté, qui ne change pas. Selon Sankoff (2005 : 1011), il faut distinguer ici deux sous-modèles.

a. la **gradation par l'âge** (*age grading*) à proprement parler : des groupes de locuteurs d'une même communauté se mettent à employer un trait linguistique, génération après génération, mais l'abandonnent ensuite au cours de leur vie. Sankoff et Blondeau (2007 : 562) la définissent comme suit : « *generation after generation, individuals change as they get older, yet the community remains stable over time* ». Ce modèle est individuel et cyclique : aucun changement ne se produit à long terme à l'échelle de l'ensemble de la communauté.

b. le **changement au cours de la vie** (*lifespan change*) : certains individus modifient leur façon de parler au cours de leur vie en adoptant un trait linguistique, qui correspond par ailleurs à une tendance générale à l'oeuvre dans l'ensemble de la communauté. Sankoff (2005 : 1011) le définit comme suit : « *individual speakers change over their lifespans in the direction of a change in progress in the rest of the community* ». Ce modèle est individuel et historique : un changement se produit à long terme à l'échelle de l'ensemble de la communauté.

3) le **changement générationnel** (*generational change*) : un groupe d'individus adopte une innovation qui est adoptée par la génération suivante. Sankoff et Blondeau (2007 : 562) le définissent comme suit : « *individual may retain their childhood patterns, with each age cohort of speakers registering an increasing use of the variant upon entering the community* ». Ce modèle est **collectif** et **historique** : un changement se produit à long terme à l'échelle de l'ensemble de la communauté.

4) le **changement collectif** (*communal change*) : jeunes et vieux adoptent simultanément un nouveau trait linguistique. Tous les membres de la communauté acquièrent ensemble la nouvelle forme (Labov 1994 : 84, Sankoff 2005 : 1010).

SYNCHRONIC PATTERN	INTERPRETATION	INDIVIDUAL	COMMUNITY
flat	1. Stability	stability	stability
regular slope with age	2a. Age grading	change	stability
regular slope with age	2b. Lifespan change	change	change
regular slope with age	3. Generational change (= apparent-time interpretation)	stability	change
flat	4. Communal change	change	change

TABLE 2. Addition of a pattern reflecting lifespan change that accompanies change at the level of the community.

Figure 81 : Les cinq configurations possibles de changement linguistique à échelle individuelle et collective (Sankoff et Blondeau 2007 : 563)

3.3. Nos résultats

Nous avons observé en section 3.3 du chapitre V, p. 368, deux phénomènes : d'une part une diminution régulière de l'usage de DEF A N, la structure canonique, en fonction de l'âge, d'autre part une augmentation régulière de l'usage de DEF N A, la structure non-canonique, en fonction de l'âge. Les syntagmes DEF N DEF A sont, quant à eux, majoritairement utilisés par les locuteurs d'âge moyen (Figure 53, p. 368 et Tableau 134, p. 629, répété ici).

Tableau 134 : Structure des SN définis produits par les participants des tests 1 et 2 en fonction de la CLASSE D'AGE

	DEF N DEF A	DEF N A	DEF A N	complexes	Total
- 15 ans	16 (10,46 %)	128 (83,66 %)	9 (5,88 %)	0 (0 %)	153 (100 %)
15-39 ans	265 (50,09 %)	117 (22,12 %)	119 (22,50 %)	28 (5,29 %)	529 (100 %)
40-59 ans	115 (42,43 %)	62 (22,88 %)	83 (30,63 %)	11 (4,06 %)	271 (100 %)
+ 60 ans	24 (14,29 %)	8 (4,76 %)	125 (74,40 %)	11 (6,55 %)	168 (100 %)
Total	420 (37,47 %)	315 (28,10 %)	336 (29,97 %)	50 (4,46 %)	1121 (100 %)

La diminution et l'augmentation présentent une forme de « pente » (*slope*) déterminée par l'âge. Elle montre qu'un changement est en cours quant à la place de l'adjectif en syntagme monodéfini, en faveur de la postposition de l'adjectif. En nous appuyant sur les modèles socio-linguistiques de Sankoff et Blondeau (2007 : 562), nous pouvons postuler trois hypothèses pour l'expliquer.

- La diminution et l'augmentation observées peuvent s'interpréter comme une **gradation par l'âge** (modèle 2a). Le changement se produirait tout au long de la vie des locuteurs, qui adopteraient une nouvelle variante à chaque nouvel âge de la vie. Le changement est cyclique, puisque les locuteurs sont « instables » au cours de leur vie, mais que le reste de la communauté est « stable ». Aucun changement à long terme ne se produit pour cette dernière.

- Elles peuvent s'interpréter comme un **changement au cours de la vie** (modèle 2b). Les locuteurs adopteraient, à un moment donné de leur vie, la nouvelle variante et rejoindraient ainsi le groupe des utilisateurs de la structure, qui constitue la future majorité. Le changement

individuel reflèterait, d'une certaine manière, le changement en cours à long terme de toute la communauté.

- Elles peuvent s'interpréter comme un **changement générationnel** (modèle 3). Des groupes de locuteurs adoptent successivement et par cohortes la nouvelle variante, rejoignant ainsi la communauté des utilisateurs de la structure, qui constitue la future majorité. La jeune génération retiendrait la variante entendue pendant l'enfance dans la bouche des parents, et entérinerait ainsi le changement, qui a lieu à l'échelle de toute la communauté.

3.4. Interprétation de nos résultats

3.4.1. DEF A N

La suite DEF A N est l'ordre des mots hérité et canonique d'après les publications scientifiques sur les langues indo-aryennes anciennes, celles sur le romani en général et sur un grand nombre de variétés particulières de romani (cf. sections 1.4.3.1, p. 138). Cela est confirmé par notre corpus albanais, puisque les locuteurs âgés emploient cette forme à une fréquence plus importante.

3.4.2. DEF N A

La suite DEF N A est apparue plus tard : comment expliquer son émergence ? Nous pensons qu'elle est due à l'influence de l'albanais, qui postpose l'adjectif de façon canonique (cf. section 2.4.1.2, p. 452). Le contact linguistique est toutefois différencié en fonction des variables sociologiques.

- EDUCATION : Le contact linguistique avec l'albanais existe depuis que les groupes actuels ont migré en Albanie, soit depuis un siècle et plus. La scolarisation, relativement récente, a approfondi ce contact car c'est l'albanais qui est enseigné et pratiqué à l'école, non le romani. Elle date du XXème siècle, et notamment du régime hoxhiste qui a tenté de la généraliser. Il n'est pas parvenu à une scolarisation totale de la population, car en ce qui concerne les Roms, elle est restée lacunaire en fonction des écoles, des familles, des individus et du genre. De Soto, Beddies et Gedeshi (2005 : 53-61) affirment en outre qu'elle a dramatiquement diminué après la chute du régime.[1] La génération de Roms qui a le plus bénéficié de la scolarisation serait donc, aujourd'hui, celle des locuteurs d'âge moyen. La diminution du temps de scolarisation va souvent de pair avec une diminution du temps d'exposition intensive à l'albanais durant l'enfance. On s'attend donc à ce que les locuteurs ayant le moins fréquenté l'école se comportent comme les locuteurs âgés (qui sont censés être plus

[1] Toutefois nous n'avons pas pu enregistrer un nombre d'enfants pour le vérifier, et tous sont ou étaient scolarisés.

conservateurs). Notre corpus confirme que les locuteurs ayant le moins d'instruction ont utilisé une majorité de DEF A N. Ils ont favorisé les syntagmes monodéfinis aux syntagmes polydéfinis, et l'antéposition plutôt que la postposition de l'adjectif. Les locuteurs qui ont bénéficié d'une instruction moyenne et supérieure ont utilisé les trois structures, avec une légère préférence pour DEF N A et DEF N DEF A. Les locuteurs qui ont bénéficié d'une instruction moyenne ont favorisé les syntagmes monodéfinis aux syntagmes polydéfinis, et la postposition plutôt que l'antéposition de l'adjectif. Ceux qui ont bénéficié d'une instruction supérieure ont employé autant les syntagmes monodéfinis que polydéfinis, donc la postposition de l'adjectif plus que l'antéposition. En résumé, plus il y a d'années de scolarité, plus il y a de polydéfinitude.

- AGE : On s'attend également à ce que les locuteurs les plus jeunes préfèrent DEF A N à DEF N A puisque leur exposition à l'albanais a duré moins longtemps, à l'instar des locuteurs moins instruits. On constate que ce n'est pas le cas dans notre corpus, puisque ce sont les jeunes qui emploient le plus DEF N A. Les locuteurs d'âge moyen emploient les trois structures, avec une légère préférence pour la polydéfinitude et pour la postposition de l'adjectif (en syntagme mono- et polydéfini). Les locuteurs jeunes et âgés ont utilisé presque exclusivement la monodéfinitude ; les jeunes ont presque exclusivement employé DEF N A tandis que les anciens ne l'ont presque jamais utilisé. Comment expliquer cette préférence inattendue pour DEF N A ? D'une part, tous nos locuteurs jeunes sont ou étaient scolarisés. D'autre part, le temps passé à la maison ne signifie pas absence d'exposition à l'albanais : il a été démontré que la télévision a un impact linguistique très important (Foulkes et Docherty 1999 : 15, Androutsopoulos 2014), de même que l'Internet et l'usage des réseaux sociaux sur les smartphones. Ces médias augmentent considérablement le temps d'exposition à l'albanais de tous les Roms, quels que soit leur genre, âge et niveau d'éducation.

- GENRE : Presque toutes les femmes de plus de quarante ans qui ont participé aux tests n'ont pas ou peu fréquenté l'école. Elles travaillent à la maison sans en sortir fréquemment ni dépasser le périmètre du quartier. Elles emploient donc l'albanais à une fréquence moindre que les hommes et dans des situations restreintes : il s'agit d'interactions occasionnelles, commerciales (ou professionnelles, pour celles qui exercent une profession hors de la maison) avec les Non-Roms. Au contraire, les hommes de plus de quarante ans qui travaillent ou ont travaillé hors de la maison connaissent un usage plus large de l'albanais, qu'ils emploient plus fréquemment et des situations plus diverses : commerciales, professionnelles, officielles comme informelles. On s'attend donc à ce que les femmes utilisent de manière préférentielle la structure conservatrice DEF A N et n'emploient pas la structure DEF N A. On constate dans notre corpus que les femmes emploient une majorité de syntagmes polydéfinis DEF N DEF A. Comment l'expliquer ? Nous avons vu en section 3.1.3, p. 361, que les femmes peuvent être porteuses d'innovation, pour peu que celle-ci soit considérée comme positive par la communauté. Ce sont alors elles qui amènent les hommes à utiliser la variante nouvelle et non-standard (Meyerhoff 2011 : 225). De plus, le genre se corrèle au niveau d'éducation : les femmes sont en moyenne moins scolarisées et moins longtemps. Elles tendront donc moins à employer la structure innovative DEF N A. Le genre et le niveau d'éduation dépendent quant à eux de l'âge. Les participantes de plus de 39 ans ont été scolarisées pendant un temps très réduit (moins de 5 années d'école primaire)

voire pas scolarisées du tout[1]. Les participantes de moins de 39 ans ont été plus nombreuses à aller à l'école, mais pas nécessairement pendant une plus longue période. Certaines ont dû quitter l'école à l'âge où elles sont devenues nubiles. On peut dire toutefois que la génération des moins de 39 ans, femmes et hommes, a acquis les bases de la lecture et de l'écriture.

3.4.3. DEF N DEF A

- GENRE : Pourquoi DEF N DEF A est-il plus utilisé par les femmes que par les hommes ? Selon nous, les femmes le préfèrent en raison de son importante extension, puisqu'il existe à la fois en grec (langue de prestige) et en albanais (langue dominante nationale), avec le déterminant adjectival portant sur les adjectifs de classe 2. Au contraire, DEF A N n'existe pas en albanais (sauf cas très marqués) et existe en grec ; DEF N A est d'un usage limité en albanais (aux quelques adjectifs de classe 1) et n'existe pas en grec. Selon Foulkes et Docherty (1999 : 16), les femmes seraient sensibles à l'extension géographique, préférant les formes non-locales aux formes locales, pour peu que les premières bénéficient d'une plus grande extension et d'un plus grand prestige linguistique que les secondes. Elles seraient ainsi plus innovatives que les hommes sur le plan syntaxique, si la forme innovative tend à refléter une variante prestigieuse. Cela peut être en rapport avec les importantes responsabilités des femmes sur le plan linguistique, dans une société où les rôles de genre sont fortement différenciés (Chambers 2013 : 306).

- AGE : Pourquoi DEF N DEF A est-il plus utilisé par les locuteurs d'âge moyen que par les jeunes et les anciens ? Nous l'expliquons par un changement en cours dans l'ordre des constituants du syntagme nominal. Si l'ordre hérité est DEF A N et que l'ordre innovant est DEF N A, DEF N DEF A représente un entre-deux, puisqu'il s'agit d'une structure très ancienne dans la langue (cf. sections 1.2, p. 408, et 1.5, p. 444), à l'instar de DEF A N… mais où l'adjectif est postposé, à l'instar de DEF N A. La génération d'âge moyen se situe elle aussi dans l'entre-deux pour appuyer cette transition.

- VARIETE : Pourquoi DEF N DEF A est-il plus utilisé par les Mečkars et les Čergars I ? Les Mečkars et les Čergars I favorisent DEF N DEF A et DEF N A mais utilisent très peu DEF A N. Le mečkar, première variété de romani présente en Albanie, a été la première à recevoir l'influence de l'albanais, langue à polydéfinitude et postposition de l'adjectif. Le čergar I est une variété *vlax* influencée du roumain, qui a pu arriver en Albanie en ayant maintenu la polydéfinitude et déjà acquis la postposition de l'adjectif. Au contraire, les Arlis ont utilisé plus de syntagmes monodéfinis que de syntagmes polydéfinis, et une majorité de DEF A N (qu'ils sont presque les seuls à employer). On sait que l'arli a été fortement influencé par le turc, langue sans article défini et à antéposition de l'adjectif, et par le

[1] A l'exception notable d'une femme mečkar de soixante-six ans, enseignante et militante pour l'alphabétisation des enfants, notamment des filles.

grec, langue à article défini et polydéfinitude mais à antéposition de l'adjectif. Ainsi doit-on diviser les locuteurs en sous-groupes qui expérimentent différents modèles de changement linguistique :

- Les **Mečkars** et les **Čergars** correspondent au modèle de **stabilité** (1), car ils ont déjà achevé le changement de l'ordre DEF A N vers DEF N DEF A puis DEF N A.
- Les **Arlis âgés** correspondent au modèle de **stabilité** (1), car ils ne sont pas « concernés » par le changement, ils n'utilisent pas de variante innovante.
- Les **Arlis d'âge moyen** ne correspondent pas exactement au modèle de gradation par l'âge (2a) car ils utilisent DEF N DEF A au lieu de la structure canonique DEF A N, qui a dû être leur variante préférentielle durant l'enfance. Ils ne correspondent pas non plus au modèle de changement générationnel (3), car la génération des Arlis jeunes ne suit pas exactement leur usage de l'adjectif. Ils correspondent plutôt au modèle du **changement au cours de la vie** (2b) : de façon individuelle, cette cohorte a changé sa façon de parler en direction du reste de la communauté que constituent les Mečkars et les Čergars.
- Les **Arlis jeunes** montrent une remarquable homogénéité. En l'absence d'étude longitudinale, il est difficile de savoir à quel modèle ils correspondent. On peut faire à leur sujet deux prédictions différentes : ils correspondent au modèle de la **gradation par l'âge** (2a) si, en grandissant/vieillissant, ils augmentent leur taux d'usage de DEF N DEF A. Ils correspondent au modèle de **changement générationnel** (3) si leurs enfants, la génération qui va leur succéder, suit leur exemple et emploie DEF N A de manière catégorique.

Selon Matras et Sakel (2007 : 848-849), tout processus de changement linguistique, induit ou non par le contact, commence au niveau individuel du locuteur, et au niveau de l'énoncé ponctuel en discours. Il est ensuite soumis à un processus de « sélection naturelle » de type évolutionniste, qui favorise ou entrave sa propagation dans tout ou partie de la communauté linguistique. Dans le cas du changement induit par le contact de langues, les locuteurs plurilingues sont les agents de cette évolution linguistique. C'est leur créativité qui trouve la solution aux problèmes pragmatiques, enrichissant ainsi l'inventaire structural de la langue. « *The motivation for this creativity is the need to perform effectively in communicative interaction while adhering, on the one hand, to the rules about the selection of clearly-identifiable phonological substance (matter) from the language that is appropriate in the particular context, while at the same time exploiting constructions that are available to the speaker in his/her entire repertoire of linguistic-communicative structures.* » Matras et Sakel (2007 : 854).

4. Scénario théorique

4.1. Pourquoi est née cette structure ?

Nous pensons que les locuteurs ne font rien en vain. Comme l'écrit Kyriakaki (2011 : 48) pour le grec, « *polydefinite constructions exist for a reason, and this is to convey certain meanings as well as reveal certain functions of the definite determiner* ».

4.1.1. L'article défini supplémentaire

L'analyse du corpus spontané nous a permis de constater que l'article défini romani albanais assume toutes les fonctions de l'article défini présentées par Boretzky (2000b : 44-54) (Tableau 13). Plus précisément, en cas de syntagme polydéfini avec adjectif antéposé, l'article défini sert un usage sémantique (*knowledge-dependent*), tandis que si l'adjectif est postposé, il sert aussi bien un usage sémantique que pragmatique (*situation-dependent*). L'analyse de l'article défini « supplémentaire » doit donc prendre en compte la **position de l'adjectif** : l'antéposition de l'adjectif va de pair avec un usage sémantique de l'article défini, tandis que la postposition permet des usages variés.

La postposition de l'adjectif (DEF N A et DEF N DEF A) permet de mettre l'adjectif sous focus, ce que ne permet pas l'antéposition. Or l'usage de l'article défini dépend du **type de focus** : s'il y a un effet de sélection, l'article défini a un usage anaphorique (discursif ou associatif), tandis que s'il y a un effet d'emphase, l'article défini a un usage situatif (direct ou abstrait). Dans le cas du focus à effet de sélection, l'article défini a une fonction de **restriction**. En effet, on a constaté que le contexte d'usage de la polydéfinitude n'est pas le même que celui de la monodéfinitude. Lorsque l'identification du référent est relativement aisée, donc lorsque la dénotation du nom comprend seulement deux entités, un syntagme monodéfini « suffit ». Si plusieurs éléments « concurrents » sont présents dans l'ensemble de la dénotation du nom, l'identification est délicate et l'adjectif seul peut ne pas suffire à marquer le contraste : un syntagme polydéfini est plus approprié.

On peut en déduire que la présence d'un article défini supplémentaire contribue à l'**identification** du référent. Nous savons que « le rôle d'un déterminant est de contribuer à la construction d'un référent » (Lebas-Fraczak 2009 : 78) : la présence de deux articles dans un syntagme avec un seul nom tête impliquera donc une construction plus complète de la référence. Le locuteur choisit d'ajouter l'article défini parce que le syntagme nominal est une expression définie incomplète (cf. section 1.1.2, p. 61), qui ne serait sinon pas valide pour un seul référent et ne permettrait pas d'assurer sans erreur l'identification. Un syntagme DEF A N ou DEF N A n'aurait pas d'autonomie référentielle suffisante dans tous les contextes. L'ajout de l'article défini sur l'adjectif contribue à compléter l'expression définie, en soutenant la capacité de catégorisation et de discrimination catégorielle de l'interlocuteur. Nous approuvons l'affirmation de Lebas-Fraczak (2009 : 94), selon laquelle « le référent est "construit" plutôt que "décrit", cette construction émanant de l'intention du locuteur et étant régulée

par l'action interprétative de l'interlocuteur. La prise en compte de l'interlocuteur dans la construction du sens, en donnant à la dimension communicative un caractère plus complet, se reflète dans les fonctions que nous attribuons aux différents déterminants ».

La présence d'un article défini devant le modifieur augmente le statut informatif de celui-ci. Précédé de l'article, il acquiert un statut « plus important » que s'il était « nu ». L'identification du référent se fait d'autant mieux que « le rôle d'un déterminant est de contribuer à la construction d'un référent » (Lebas-Fraczak 2009 : 6). Cela explique les effets d'emphase et de sélection des structures DEF N DEF A, mais également la quantité plus importante de structures DEF N DEF A par rapport aux structures DEF A DEF N. Les premières sont des cas de modification indirecte, les deuxièmes de modification directe. La focalisation n'est pas un phénomène binaire, où le trait [+ foc], focalisé, s'opposerait au trait [- foc], non focalisé ou défocalisé. Nous pensons qu'elle se représente mieux par un **continuum**, d'où l'hypothèse de la Figure 82, où - *foc* ne représente pas le pôle négatif [- foc] mais la structure la moins focalisée de toutes, et + *foc* la structure la plus focalisée de toutes. Il n'existe donc pas de séparation binaire entre expression définies complètes et incomplètes, mais une gradation des unes aux autres.

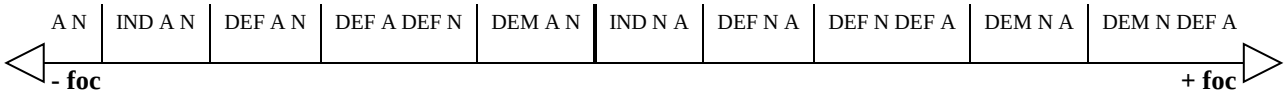


Figure 82 : Continuum de focalisation sur l'adjectif épithète dans le SN romani

Les structures avec un démonstratif sont les plus focalisées car, comme en français, « le déterminant démonstratif indique […] que le référent est à la fois présupposé et particularisé. Ainsi, son emploi permet de mieux focaliser le référent que l'emploi de l'article défini » (Lebas-Fraczak 2009 : 94). Parmi les syntagmes introduits par un article défini, ceux polydéfinis (DEF N DEF A) sont donc les plus focalisants. Le deuxième article sert à focaliser l'adjectif, par ailleurs postposé, donc prêt à recevoir le focus interne du syntagme nominal. L'adjectif reçoit ainsi une double focalisation : par sa position syntaxique et par l'ajout de matériel grammatical. Creissels (2006 : 126) constate qu'il existe « des langues où les mécanismes de focalisation combinent déplacement du terme focalisé et adjonction d'un marqueur de focalisation » : le romani en fait partie. En ce sens, la place de l'adjectif aussi bien que le déterminant supplémentaires contribuent au marquage de la focalisation dans le syntagme nominal (Lebas-Fraczak 2017 : 75).

L'article défini supplémentaire joue donc le rôle d'**opérateur d'identification ou de restriction** (cf. Manolessou 2000). Cela explique pourquoi le syntagme polydéfini donne l'impression d'être plus spécifique qu'un syntagme monodéfini. Il n'y aurait pas à proprement parler de *polydéfinitude*, en ce sens qu'il n'y a pas nécessairement plusieurs instances de l'article défini (mais plusieurs instances de déterminant défini : POSS N DEF A, DEM N DEF A). Il n'y a pas nécessairement non plus de *poly-définitude* car le syntagme n'est pas sémantiquement plus défini, le deuxième article défini n'a pas pour fonction de rendre plus défini le syntagme nominal. Nous choisissons donc de dissocier article défini et définitude, comme le proposent Lekakou et Szendrői (2012 : 116). Le romani présenterait donc la même caractéristique que le grec.

Nous considérons que l'article « adjectival » n'est pas un « véritable » déterminant, mais un **morphème focalisateur**. Il ne présente pas les caractéristiques habituelles du déterminant (cf. section 1.1.1, p. 59) car il ne satisfait pas les contraintes d'unicité et d'existence (Charolles 2002) : l'adjectif n'est pas concerné par ces catégories. Il ne satisfait pas non plus l'éventuelle contrainte de familiarité, car le statut informatif de l'adjectif est en général nouveau. Lebas-Fraczak (2015 : 73) considère l'article (défini, indéfini, partitif) du français comme un « focalisateur » du nom, et l'absence d'article comme « un indice du statut défocalisé du nom » (Lebas-Fraczak 2015 : 73) « afin de signaler que ce n'est pas cet élément qui doit focaliser l'attention de l'interlocuteur en particulier » (Lebas-Fraczak 2017 : 78). Nous pouvons postuler la même hypothèse pour l'article défini des syntagmes polydéfinis en romani, où l'article défini serait un focalisateur de l'adjectif.

Enfin, on pourrait se demander pourquoi c'est l'article défini qui assume cette fonction, et non un autre élément de la langue qui servirait seulement à la focalisation. C'est le cas des particules de focalisation *em, i, ta,* et, *i li, ta li,* et même, *ja i li / ja li,* oui même, *li,* même, *vi,* aussi, etc. On peut supposer que les locuteurs ont préféré éviter cette stratégie consistant à introduire dans le syntagme un mot « nouveau », réellement supplémentaire, mais répliquer un mot déjà présent. Cette explication synchronique serait donc le principe d'économie. De plus, tous les articles n'ont pas la même **force focalisatrice**. Selon Lebas-Fraczak (2017 : 78) et dans le cadre de l'approche interlocutive « l'emploi de l'article défini se base sur la présupposition d'une connaissance chez l'interlocuteur relativement à la "chose signifiée" correspondante au SN, alors que les articles indéfini et partitif marquent son statut non présupposé, nécessitant de mobiliser plus fortement l'attention de l'interlocuteur ». Dans le contexte d'emploi qui nous intéresse, l'article indéfini, non présupposant, n'aurait pas de sens.

Un dernier argument va dans le sens de la focalisation. On sait que plus une expression est longue ou « lourde », plus elle participe de la focalisation. La quantité d'expressions qui forment un élément est symptomatique du statut communicatif des éléments au sein d'un énoncé : « plus elle est grande plus l'élément est focalisé, un élément omis, voire exprimé « brièvement » (pronominalisé), correspondant à une connaissance présupposée et étant défocalisé » (Lebas-Fraczak 2015 : 32). Or ajouter un article défini dans le syntagme consiste bien à l'allonger, mais aussi à l'alourdir sur le plan grammatical.

4.1.2. La place de l'adjectif

Notre étude contredit l'hypothèse selon laquelle la présence de l'article défini après le nom s'explique uniquement par la présence de l'adjectif après le nom, qui ne peut pas occuper seul cette position (Matras 2000b : 103-104). Nous avons constaté que DEF N A et DEM N A existent ; la portée du déterminant qui introduit le syntagme inclut bel et bien l'adjectif postposé, malgré l'affirmation de Matras (2000b : 103) en ce qui concerne le déterminant démonstratif : « *One could hence argue that the postpositioned adjective falls outside the determiner scope of the demonstrative, and demands its own additional determination* ».

De plus, les exemples étudiés montrent que l'emploi différencié de l'adjectif n'est pas qu'une question de place canonique/non-marquée ou non-canonique/marquée. Il s'agit de manières différentes de présenter l'information.

- La place antéposée permet un emploi moins spécifique voire générique de l'adjectif – d'où un emploi pour délivrer une propriété inhérente au nom, pour référer de manière imprécise ou hyperonymique. L'adjectif fait partie intégrante de l'expression référentielle globale.
- La place postposée permet un emploi spécifique, discriminant, classificateur. Il permet d'identifier et de pointer le référent. La référence de l'expression globale se trouve à l'intersection des référents du nom et de l'adjectif.

Le romani présente la même topologie syntaxique et sémantique que les langues romanes selon Cinque (2010) et Cinque (2014) : la position antéposée est celle de la modification directe, tandis que la position postposée est celle de la modification directe ou indirecte.

La postposition de l'adjectif, qu'il soit articulé ou non, a plusieurs fonctions. Nous avons fait l'hypothèse, représentée par la Figure 33 p. 250, qu'une partie des emplois de DEF A N a été, à un moment donné de l'histoire, assignée à une nouvelle construction, DEF N A. Il se serait produit une redistribution des fonctions de l'adjectif épithète en fontion de sa place dans le syntagme nominal. Une première fonction de DEF N A est, pourrait-on dire, l'**« accueil » des adjectifs empruntés**. Nous avons observé en section 2.4.3 du chapitre V, p. 335, que l'emplacement postnominal est préféré pour les adjectif empruntés – qu'ils l'aient été à une langue antéposant ou postposant l'adjectif, qu'ils aient été morphologiquement adaptés ou non, et quelle que soit l'époque de l'emprunt. En ce sens, le romani contemporain a compartimenté sa syntaxe nominale.

La deuxième est de permettre un **focus étroit sur l'adjectif**. Selon nous, le syntagme nominal en romani d'Albanie correspond au modèle proposé par Nølke (2001 : 180) pour le français :

> « Antéposition : (Adj (Sub))
> Postposition : (Sub (Adj))
>
> Découlent directement de ces structurations les deux règles suivantes :
> - Un adjectif antéposé ne constitue jamais à lui seul le foyer simple : ou bien il est focalisé avec son substantif, ou bien il se trouve en dehors du foyer.
> - Un adjectif postposé est toujours focalisé : ou bien il forme le foyer avec son substantif, ou bien il est seul à être focalisé. » (Nølke 2001 : 180)

En construisant son syntagme nominal, le locuteur effectue un choix entre l'antéposition et la postposition de l'adjectif selon une certaine « intention » qui est la « visée de la focalisation » (Nølke 2001 : 92). Le placement de l'adjectif est donc libre mais régi par des contraintes préférentielles et pragmatiques. Cette théorie implique que « quelque chose est toujours focalisé au sein d'un SN avec un adjectif : ou bien le SN entier (ce qui peut être le cas indifféremment de la position de l'adjectif) ou bien le nom seul (en cas d'antéposition de l'adjectif) ou bien l'adjectif seul (en cas de postposition de l'adjectif) » (Lebas-Fraczak 2017).

Nous avons observé, avec le corpus spontané, que l'adjectif postposé transmettait le plus souvent une **information nouvelle**, ce qui va dans le sens de l'affirmation de Lebas-Fraczak (2015 : 34) : « la

place initiale [est] souvent occupée par un élément déjà connu de l'interlocuteur, alors que ce qui suit constitue une information nouvelle, celle qui possède donc une plus grande importance communicative. Ainsi, la tendance générale pour tout énoncé est d'aller du connu vers le nouveau. »

On pourrait s'étonner qu'une théorie de la focalisation élaborée pour le français soit pertinente pour le romani, qui n'est pas une langue romane. Selon cette théorie, c'est la place finale du syntagme qui est focalisante, « à la différence de la place initiale, laquelle est souvent liée à la présupposition » (Lebas-Fraczak 2015 : 31). Mais (Nølke 2001 : 180) fait le lien entre emplacement du focus et accent prosodique :

> « La focalisation simple est liée directement à la présence de montées et de descentes mélodiques qui, précisément, fonctionnent comme délimitateurs de groupes rythmiques. […] Il découle de ces définitions qu'une focalisation a lieu à l'intérieur de chaque syntagme substantival. En effet, la fin d'un tel syntagme est toujours marquée par un délimitateur prosodique. […] En d'autres termes, chaque syntagme substantival constitue un domaine de focalisation mineure. » (Nølke 2001 : 180)

Or le français présente une caractéristique qui le distingue des autres langues romanes : il ne possède pas d'accent de mot ; l'accent ne joue aucun rôle distinctif permettant de distinguer des homonymes dans des paires minimales. Les mots n'ont, « dans la chaîne parlée, aucune identité sonore perceptible : l'unité de rang immédiatement supérieure à la syllabe est le groupe accentuel (ou groupe rythmique) » (Riegel, Pellat et Rioul 2009 : 57). L'accent « est tout entier au service de la démarcation des groupes » (Riegel, Pellat et Rioul 2009 : 57-58). L'énoncé est alors rythmé par « la démarcation du discours en groupes accentuels, d'où le nom de groupes rythmiques qui leur est donné » (Riegel, Pellat et Rioul 2009 : 59). C'est la dernière syllabe du groupe qui porte l'accent (Riegel, Pellat et Rioul 2009 : 57). Le romani présente, lui, un accent de mot… et cet accent est placé sur la dernière[1] syllabe de chaque mot.[2] Ce fait pourrait avoir favorisé l'acquisition de la structure polydéfinie.

Le romani présente d'ailleurs une différence avec le français. Aucun adjectif n'est *a priori* agrammatical dans une position ou l'autre, comme peuvent l'être les adjectifs relationnels du français, toujours postposés (Nølke 2001 : 190-191) : *une ville danoise*, mais **une danoise ville*. Nous trouvons par exemple l'adjectif « romani » antéposé dans l'exemple (172) mais postposé dans les exemples (186), (187) et (203). Ce n'est donc pas tant l'adjectif lui-même que sa position, qui met en valeur certains sèmes plutôt que d'autres. La syntaxe nominale du romani est plus libre que celle du français.

L'adjectif, articulé ou non, n'est pas le seul élément à pouvoir être postposé. Peuvent également être postposés :

- **le déterminant possessif** : *i čhib amari*, notre langue (*M23M, Llakatund, 13 août 2014*), *i istorija amari le Romengi*, notre histoire des Roms (*M68A, Voskopojë, 15 juillet 2013*), *duj phrala vavera mire*, deux autres de mes frères (*M30M, Voskopojë, 15 juillet 2013*)
- **le déterminant démonstratif** : *i patata odəja*, cette pomme de terre (*F47M, Llakatund, 13 août 2014*), *o manykyri akola*, ce vernis à ongle (*M45M, Llakatund, 13 août 2014*), *sea-l boàle kadala*,

[1] La seule exception est la flexion nominale de la strate II : ces morphèmes agglutinants sont placés en fin de mot mais ne peuvent porter l'accent, à l'instar d'enclitiques. L'accent d'un nom fléchi à la strate II est donc paroxytonique.
[2] C'est toutefois aussi le cas du turc et de l'arménien, qui ne présentent pas pour autant de structure polydéfinie.

toutes ces maladies, *le kurùza le kodola*, ces céréales, *ël magàrea le kodola*, ces chevaux (Cioabă 2006a)

- **le déterminant numéral** : *me phen o do cignen*, mes deux petites sœurs (RMS, dialecte spoitari RO-016)
- **le nom non articulé en apposition étroite** : *amare organizata Amari Phen*, notre organisation Notre Sœur (corpus sponané de l'exemple (188), p. 595), *i čhaj Jylja*, la fille Jylja (*M27M, Llakatund, 13 août 2014*)
- **le nom articulé en apposition étroite** : *o gulipe i tatlija*, la pâtisserie tatli (*F53A, Korçë, 25 juillet 2014*), *i Jylja i čhaj*, la fille Jylja (*M27M, Llakatund, 13 août 2014*)
- **le syntagme génitival, de forme courte ou longue** : *i istorija amari le Romengi*, notre histoire des Roms (*M68A, Voskopojë, 15 juillet 2013*), *i tradicja e Romengi*, la tradition des Roms (*M68A, Voskopojë, 15 juillet 2013*), *o anav e thanesko kaj beshel*, le nom du lieu où il habite (*M50A, Voskopojë, 15 juillet 2013*), *kryetar e Albanijakoro*, président de l'Albanie (*M52A, Korçë, 25 juillet 2014*)

Dans la tradition qualifiée par Verhoeven et Skopeteas (2015 : 1) de « *form-to-function association between focus and constituent structure* », nous pensons que le focus interne au syntagme nominal s'exprime via une position structurellement identifiable. La structure de l'information du syntagme nominal réside dans la syntaxe. Ainsi proposons-nous un schéma « géographique » du syntagme nominal en romani, en Figure 83.

$$\underbrace{\text{DEF}_1 \quad \text{A}}_{\text{-foc}} \quad \text{N} \quad \underbrace{\text{DEF}_2 \quad \text{A}}_{\text{+foc}}$$

Figure 83 : Topologie du syntagme nominal en romani d'Albanie

Nous avons choisi de représenter ici la suite syntagmatique DEF A N DEF A. Mais, de plus :

- à l'emplacement de DEF_1 peuvent également se trouver DEM, POSS, NUM, QUANT, IND, associés ou non à un article défini.
- à l'emplacement de DEF_1 A peut se trouver un syntagme génitival.
- à l'emplacement de DEF_2 peuvent se trouver DEM, POSS, NUM associés ou non à un article défini, ainsi qu'un nom (commun ou propre) nu ou un syntagme génitival sans article défini.
- à l'emplacement de DEF_2 A peuvent se trouver un syntagme génitival avec article défini ou un nom (commun ou propre) articulé en apposition nominale étroite.

4.1.3. Les fonctions

Nous pensons que la langue et la grammaire sont le produit de la création humaine et des outils socioculturels au service de la communication. Ainsi que l'écrit Chambers (2013 : 320-321), « *we express who we are with fine nuance and no little grace, varying the way we sound and the words we use depending upon the social setting in which we are speaking. And it is not only our own class, sex, age, ethnicity, style, and other attributes that we are sensitive to but also all those things in the people we are speaking to. The social uses of language are so fine-tuned and subtle as to be virtuosic, but they are mostly beyond conscious manipulation or simple introspection* ». Nous pensons également que le changement linguistique est motivé par l'usage qu'ont les locuteurs de leur langue, et que le changement syntaxique est causé par des raisons sémantiques et pragmatiques (Waltereit et Detges 2008 : 14).

Nous connaissons la fonction de l'article défini supplémentaire et celle de la place de l'adjectif : tentons maintenant de comprendre celle de l'ensemble de la sous-partie adjectivale du syntagme polydéfini. L'article supplémentaire comme la postposition de l'adjectif ne sont pas requis par la grammaire de la langue, mais font l'objet d'un choix du locuteur parce qu'ils jouent un rôle pragmatique, et exercent des fonctions utiles lors de la construction et de la communication de l'énoncé. Nous avons pu relever les fonctions suivantes :

- **contraste et sélection** : c'est le cas des exemples (154) à (161), présentés en section 2.1.1.1, p. 211. C'est également le cas de toutes les occurrences produites lors de la passation des tests, conçus pour provoquer l'usage de syntagmes polydéfinis revêtant cette fonction.

- **emphase sur l'adjectif** : c'est le cas des exemples (162) et (163), présentés en section 2.1.1.2, p. 218.

- **précision sur un référent donné** en discours : c'est le cas des exemples (158), (159), (160), (165), (166) ou (170) du corpus spontané. On observe que le nom est alors répété, à moins qu'il fasse place à un synonyme, un hyponyme ou un hyperonyme. Voici les exemples (159) et (170) :

(159) *F38A, Korçë, 4 août 2014*

Akəna	*al-i*	*i*	*čhaj*	*kat-o*	*kampi/*
maintenant	venir.PRET-3SG.F	ART.DEF.F.SG	fille(M.SG)	venant_de-ART.DEF.M.SG	camp(M.SG)
i	***čhaj***	***i***	***vaver.***		
ART.DEF.F.SG	fille(M.SG)	ART.DEF.F.SG	autre(F.SG)		

« La fille vient de rentrer du camp / **l'autre fille**. »

(170) *M70/71A, Korçë, 28 juillet 2014*

Na	*mukh-l-e*	*la*	*o*	*čhaj-a/*	***o***	***tikn-e***	***o***	***maksemi.***
NEG	laisser-PRET-PL	3.SG.F.OBL	ART.DEF.PL	fille-F.PL	ART.DEF.PL	petit-PL	ART.DEF.PL	enfant(M.SG)

« Les filles, **les petits enfants** ne l'ont pas laissée [partir]. »

Cette fonction s'observe également en romani de Roumanie (Cioabă 2006a) :

(276) Romani, Cioabă (2006a)

Del o brïšind, thowaw iw thaj kaθa' jekh morći guruvesqi p-o punro cërùlea, kë beśas anθ-o drom e reat le iwendesqi. ***E reat e bari*** *le iwendesqi.*

« Il pleut, je jette la neige d'un morceau de cuir sur mon pied bleui, parce que nous nous trouvons sur la route par la nuit d'hiver. **La longue nuit** de l'hiver. »

- **synthèse du discours qui précède** : c'est le cas de l'exemple (169) du corpus spontané. La locutrice vient de dépeindre la misère dans laquelle elle vit avec ses enfants : disette, manque de vêtements, absence de chauffage, logement insalubre, pas d'électricité, impossibilité de scolariser les enfants… Elle conclut que le pire des problèmes est la faim, notamment pour son bébé qui ne peut endurer la privation de nourriture.

(169) *F24A, Bilisht, 27 juillet 2014*

O	*xabe*	*si*	***o***	***problemi***	***o***	***bar-o.***
ART.DEF.M.SG	nourriture(M.SG)	être.3.PRES	ART.DEF.M.SG	problème(M.SG)	ART.DEF.M.SG	grand-M.SG

« La nourriture c'est **le grand problème**. »

- **surnom** ou « étiquette » (*epithet*) : c'est le cas de l'exemple (171) du corpus spontané.

(171) *[Zamira], Korçë, 1er août 2014*

O	***barval-o***	***o***	***Skender***	*so*	*ker-d-as?*
ART.DEF.M.SG	riche-M.SG	ART.DEF.M.SG	Skender(M.SG)	INT	faire-PRET-3SG

« **Skender le riche** qu'est-ce qu'il a fait ? »

Un informateur mečkar m'a moi-même surnommée « la chatte noire » lors d'un échange écrit sur Internet[1] :

(277) *M30M, 6 juin 2014*

Hey	*so*	*Ker-es*	*tu*	***I***	***Macka***	***I***	***Kal-i***
hey	quoi	faire-2SG	2SG	ART.DEF.F.SG	chatte(F.SG)	ART.DEF.F.SG	noir-F.SG

« Hey, comment ça va, **Chatte Noire** ? »

Mon hôte de Korçë m'a, quant à lui, surnommée « la chatte blanche » (peut-être en référence au film *Chat noir, chat blanc* d'Emir Kusturica) :

(278) *M46A, Korçë, 24 juillet 2014*

Tu	*isin-an*	***i***	***mačka***	***i***	***parn-i***
2SG	être-2SG	ART.DEF.F.SG	chatte(F.SG)	ART.DEF.F.SG	blanc-F.SG
i	*mačka*	*e*	*Franciaz-a-q-i*		
ART.DEF.F.SG	chatte(F.SG)	ART.DEF.F.OBL	France-F.OBL-GEN-F.SG		

« Toi tu es **la chatte blanche**, la chatte de France. »

[1] Nous respectons ici son orthographe.

Cette fonction s'observe également en romani de Roumanie. Le corpus de Cioabă (2006a) mentionne différentes personnes portant des surnoms : *o Maco o phuro*, Maco le vieux, *o Parolea o phuro*, Parolea le vieux, *o Iono o Rëguśime*, Ion l'enroué (à la voix rauque)…

Enfin, cette fonction est utilisée pour traduire des noms propres en romani. C'est le cas du personnage *Dănilă Prepeleac*, Danilo le bêta, dans le conte éponyme de Ion Creangă. Pour le manuel scolaire de langue et littérature romani (Sarău et Cordovan 2011 : 33-40), Lidia Ursu l'a traduit par *o Dänìlä o Dilo*, Danilo le bêta/le fou, ou encore par *o Dänìlä o Ćakado*, Danilo l'aveugle.

- **« romanisation » d'un terme** ou d'un référent étranger : c'est le cas de l'exemple ci-dessus, où un (sur)nom propre roumain présentant un adjectif épithète est traduit en romani à l'aide d'un syntagme polydéfini. C'est également le cas de l'exemple ci-dessous, où le terme *temberon*, poubelle, est emprunté au roumain et où le mot « poubelle noire » forme un quasi mot composé dans le contexte. Ici, la structure polydéfinie sert à rendre l'expression empruntée plus « acceptable » en romani, à favoriser son insertion dans l'énoncé – ce que nous nommons ici ponctuellement « romanisation ».

(279) Romani, Ţăndărei (Roumanie)
*Put other rubbish in **the black bin**.*
*Pui gunoiul menajer în **tomberonul negru**.*
*Ciu sea okaver o gunoi and**o temberon o kalo**.*
« Jetez le reste des déchets dans **la poubelle noire**. »

Cet exemple provient d'un prospectus trilingue (anglais, roumain, romani) visant à inciter les habitants des quartiers rom de Manchester à pratiquer le tri sélectif. Il a été conçu par l'équipe du projet Migrom de l'université de Manchester en 2014-2015, et traduit[1] par une participante du projet, Rom roumaine originaire de Ţăndărei.

4.1.4. L'optionnalité

Selon Thuilier (2015 : 271), « la connaissance langagière des locuteurs n'est pas définie uniquement en termes de contraintes catégoriques, mais […] elle embrasse également des préférences ». Elle examine les *contraintes préférentielles* qui interviennent dans « des phénomènes non catégoriques pour lesquels il existe plusieurs réalisations concurrentes » (Thuilier 2015 : 271). Dans de telles situations grammaticales, les locuteurs doivent opérer un choix pour exprimer au mieux leurs idées tout en tenant compte de leur interlocuteurs et de la situation d'énonciation. Ces contraintes interviennent à tous les niveaux linguistiques, la prosodie, la phonologie, la syntaxe, la sémantique comme la structure informationnelle. Dans le cas de la syntaxe et en particulier de l'ordre des constituants nominaux, plusieurs facteurs entrent en ligne de compte : la catégorie des éléments et les contraintes qui s'y appliquent (par exemple, le déterminant précède le nom en romani), les contraintes structurelles (par exemple, tout ce qui n'est pas la tête précède la tête, selon l'ordre hérité du romani),

[1] Nous respectons ici son orthographe.

les traits (par exemple le poids) (Thuilier 2015 : 279). Nous pensons que la polydéfinitude fait l'objet de contraintes préférentielles en romani.

C'est pourquoi elle est très proche du redoublement syntaxique (*syntactic doubling*) (Barbiers 2013, Barbiers, Sjef, Koeneman et Lekakou 2009) mais n'en est pas un. Le redoublement syntaxique est un phénomène de variation syntaxique où « a morphosyntactic feature, morpheme, word or phrase is redundantly expressed two or more times within a sentence » (Barbiers 2013 : 6). Barbiers (2013 : 9-10) dresse une typologie des phénomènes de redoublement, résumés par le Tableau 111.

Tableau 111 : Les différents types de redoublement syntaxique Barbiers (2013 : 9-10)

total	régional	optionnel	fonctionnel + fonctionnel	proche	apparent
partiel	familier	obligatoire	fonctionnel + lexical	distant	réel
			lexical + lexical		

La polydéfinitude en romani d'Albanie ressemble fort à un cas de redoublement total, régional, optionnel, impliquant deux morphèmes fonctionnels, proche (ou local) et réel. Toutefois, la définition exacte du *syntactic doubling* est la suivante : « *A construction involves syntactic doubling (i) if a sentence* S_A *in language variety A is semantically and pragmatically equivalent to sentence* S_B *expresses it only once, or (ii) if within a language variety a constituent can optionally be doubled without any detectable effect on the semantic or pragmatic interpretation of that sentence* » (Barbiers 2013 : 7). La polydéfinitude du romani ressemble au type (ii) où le morphème ART.DEF est copié à l'identique… mais nous avons bel et bien détecté un effet et une visée pragmatique au redoublement. Quand bien même il est, par exemple, difficile voire impossible de traduire dans une langue étrangère un syntagme polydéfini d'une manière différente d'un syntagme monodéfini, la différence existe bien, en ce qui concerne l'usage qu'en ont les locuteurs. Elle a un effet sur la sémantique et la pragmatique de la phrase. Ce n'est donc pas du redoublement syntaxique.

Barbiers (2013 : 7) considère qu'il n'y a variation que si deux syntaxes différentes correspondent à la même interprétation et que si la redondance ne contribue nullement à l'interprétation sémantique ou pragmatique. Nous pensons au contraire qu'en romani il peut y avoir variation syntaxique et contribution pragmatique, donc que la polydéfinitude relève bien de la variation. Un locuteur peut choisir d'employer DEF N A même dans un cas où DEF N DEF A serait plus adéquat et rendrait plus aisée l'identification pour l'interlocuteur. Cela reste une option. Gardiner (1989), reconnaissant l'importance de l'interlocuteur dans l'acte de parole, écrit qu'« il faut que le locuteur se préoccupe de la manière dont son partenaire pensera le référent pour que la pensée puisse être communiquée » (Gardiner 1989 : 230). Le locuteur élabore pour cela des stratégies de mise en relief (Lebas-Fraczak 2017 : 76). La polydéfinitude romani en est une et montre que celles-ci sont affaire de choix.

On sait que la relation interlocutive s'exerce dans les deux sens (Lebas-Fraczak 2017 : 76). L'interlocuteur influence le locuteur qui prend en compte les connaissances supposées de son interlocuteur (c'est la *présupposition*) lors de l'acte de communication. Le locuteur cherche à influencer son interlocuteur en orientant son attention et son interprétation (grâce à la *focalisation*). Faut-il voir dans les cas d'emploi de DEF N A « au lieu de » DEF N DEF A des formes relatives d'échec

de la relation interlocutive ? En effet, on sait que l'emploi d'une expression définie incomplète révèle quelle capacité d'interprétation le locuteur attribue à son interlocuteur, et révèle qu'il suppose son interlocuteur capable d'interpréter le syntagme, de repérer le référent comme unique et de l'identifier sans difficulté (section 1.1, p. 59). L'usage de la polydéfinitude marque un souci de rendre plus « complète » l'expression définie. Dans ce cas, l'emploi de la monodéfinitude révèle que le locuteur estime ou surestime les capacités d'interprétation de l'interlocuteur, en le laissant « se débrouiller ». Le locuteur peut également n'être pas conscient des difficultés d'interprétation et d'identification, ou de la présence de référents « concurrents » dans la dénotation du nom.

4.2. Le processus de transfert linguistique

4.2.1. Le mécanisme de changement syntaxique

La modification des relations syntaxiques et de l'ordre des mots au contact d'une langue étrangère est un cas de transfert linguistique (Heine 2008 : 36). On peut en effet transférer plusieurs éléments linguistique : la forme (les sons), le sens (lexical ou grammatical), les relations syntaxiques ou une combinaison de plusieurs de ces éléments (Heine 2008 : 36). Il existe deux types principaux de transfert linguistique : l'emprunt et la réplication. Le premier implique un transfert direct de matériau linguistique (mot, phonème...) depuis une langue étrangère. Le deuxième n'implique rien de tel : les locuteurs utilisent les moyens disponibles dans leur langue afin de créer de nouveaux mots ou mécanismes grammaticaux, les plus similaires possible à ceux de la langue étrangère (Heine 2008 : 57). La Figure 84 schématise les principaux types de transfert linguistique induits par le contact.

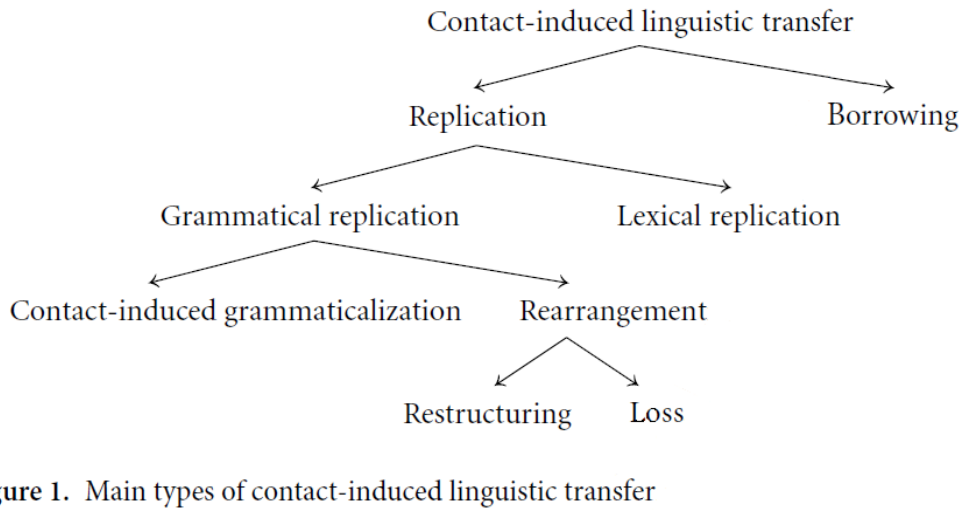


Figure 84 : Les principaux types de transfert linguistique induits par le contact (Heine 2008 : 37)

La réplication peut être de deux types : lexicale (qui concerne le vocabulaire) ou grammaticale (qui concerne les structures) (Heine 2008 : 37). On entend par « réplication grammaticale » ce qui se nomme ailleurs « calque » ou « emprunt structurel » (Heine 2008 : 36). Elle consiste à créer un nouveau sens ou une nouvelle structure en utilisant les ressources linguistiques propres à une langue, mais sur le modèle de la structure d'une langue étrangère. La langue d'accueil, cible ou receveuse, se nomme langue réplique (R). La langue étrangère, langue source ou donneuse, se nomme langue modèle (M) (Heine 2008 : 36, Heine et Kuteva 2005).

Il existe quatre processus menant à un changement d'ordre des mots : « *(a) narrowing (of options), (b) shift from one construction type to another, (c) pragmatic unmarking, and (d) extension and frequency* » (Heine 2008 : 38). Les stratégies sont les suivantes :

> « *a. Select among the word order alternatives that exist one that matches the order to be found in the model language.*
> *b. Use an existing construction and assign it a new function.*
> *c. Use a pragmatically marked construction and develop it into a pragmatically unmarked construction.*
> *d. Extend an existing use pattern to new contexts.*
> *e. Use an existing use pattern more frequently.* » (Heine 2008 : 56)

Le plus souvent, ces stratégies se combinent car elles sont toutes des manifestations différentes d'un même processus général (Heine 2008 : 57). Dans le cas qui nous intéresse, celui du changement des relations syntaxiques, il est plus économique, pour les locuteurs d'une langue R, de « recruter » une structure déjà existante dans leur langue et de la redéfinir, afin de créer une structure qui reflète l'ordre des mots de la langue M (Heine 2008 : 56). La « nouvelle » structure de la langue R n'est donc pas à proprement parler nouvelle, puisqu'elle est construite sur du matériau existant et déjà disponible dans la langue. (Heine 2008 : 57) note que, même si le changement opéré est d'ordre syntaxique, la stratégie employée est plutôt sémantique ou pragmatique. Le contact de langues est un facteur déclencheur et/ou catalyseur.

Dans le cas qui nous intéresse, on peut également parler de convergence. On appelle « convergence » le phénomène de réorganisation syntaxique sur le modèle d'une langue étrangère. Elle présente à la fois le caractère d'une rétention et d'une accommodation. Matras (2000a : 190) et Matras et Sakel (2007 : 835) la définissent comme « *the adaptation of an internal element to match the structure and scope of an external element in the contact language, which is perceived as its functional counterpart* ». Cette notion est proche de celle présentée ci-dessus, mais elle assigne également un objectif au phénomène de transfert linguistique. Le processus de convergence se fait en trois étapes Matras (2000a : 190-191) :

- identifier des constructions syntaxiques qui ont les mêmes fonctions dans les langues R et M,
- identifier les pivots fonctionnels qui sont au fondement de ces constructions, dans chacune des deux langues,
- identifier les phénomènes de variation paradigmatique applicables aux pivots, comme la variation dans l'ordre des mots,

- identifier les traits de la langue R qui, au contraire, ne souffrent pas de variation et de compromis,
- sélectionner les contraintes et effectuer la « greffe ».

L'étape d'identification des phénomènes de variation est celle qui permet de créer un compromis grammaticalement acceptable pour la langue R. Ce compromis autorise le changement structural, l'adaptation interne sur la base d'un modèle externe.

Le processus de convergence se produit toujours avec un objectif communicatif donné. « *Pattern replication is a functional phenomenon in the first instance, one that differs from matter replication or MAT* [*matter replication*] *in taking into account the social constraints on importation of matter* » Matras et Sakel (2007 : 858). La raison d'une telle innovation linguistique est la volonté, de la part des locuteurs, d'adopter une solution pragmatique aux problèmes qui se posent dans la structuration de leurs énoncés. Comme il s'agit d'un phénomène induit par le contact de langue, la convergence ne peut se produire qu'en contexte socio-linguistique bilingue ou multilingue, le seul à même de valider et de régulariser la solution pragmatique innovante, et ainsi de mener au changement linguistique (Matras et Sakel 2007 : 857). Notons qu'il peut s'agir d'un ajout à l'inventaire structurel de la langue comme d'une suppression[1]. La *pattern replication* est donc exclue lorsqu'aucune structure n'est disponible dans la langue R pour assumer le rôle de trait pivot de la construction dans la langue M (Matras et Sakel 2007 : 847). La Figure 85 schématise les processus de *pattern replication* et de *pivot-matching* (Matras 2009 : 242).

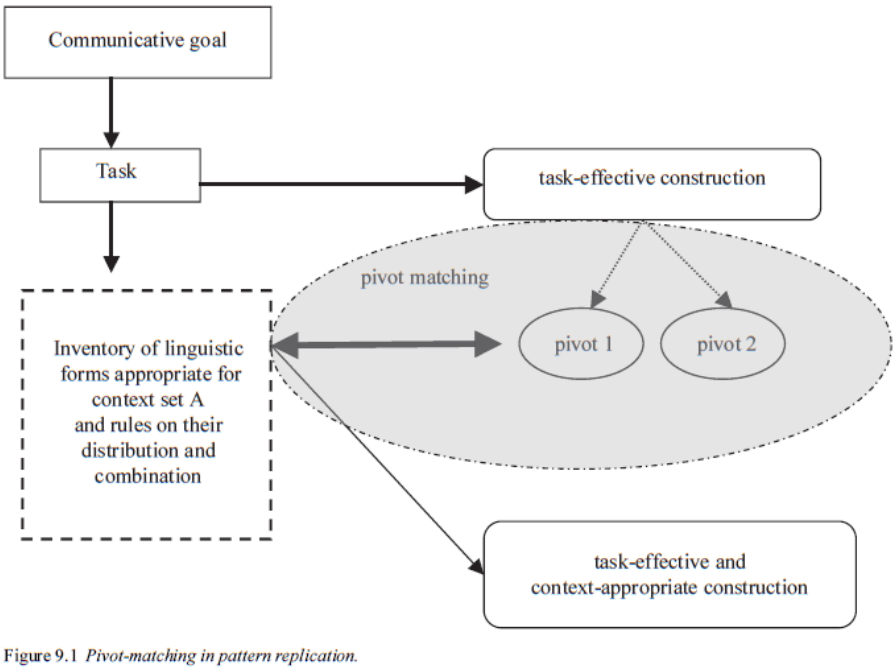


Figure 85 : Le processus de *pattern replication* et de *pivot-matching* (Matras 2009 : 242)

[1] C'est le cas des variétés de romani en contact avec le polonais et le russe, qui ont perdu l'article défini et même l'article indéfini (Matras 2002 : 96). Une *pattern replication* a eu lieu selon Matras et Sakel (2007 : 859), où le trait pivot est le marquage de la définitude sur le nom, de façon contextuelle plutôt que formelle.

On voit que la convergence permet d'opérer des compromis entre deux systèmes linguistiques. Matras et Sakel (2007) et Matras (2009 : chap. 9) distinguent deux types de convergence : d'une part la réplication directe de formes phonétiques et de morphèmes depuis la langue source, qu'ils nomment *matter replication* (réplication de matériau), et d'autre part la *pattern replication* (réplication de motif), qui est la réplication de configuration syntaxique, de distribution, de sens grammatical, d'arrangements formels au niveau du discours, de la phrase, du syntagme ou du mot. Les deux types de réplication s'excluent souvent l'un l'autre (Matras et Sakel 2007 : 830). La *pattern replication* se produit suite aux étapes présentées ci-dessus, et notamment la recherche et l'accommodation de pivots (*pivot-matching*) dans les deux langues concernées.

Ce modèle théorique s'applique tout particulièrement au phénomène qui nous intéresse, car certaines propriétés structurelles, si elles sont empruntées, ne peuvent par nature faire l'objet que d'une *pattern replication* (Matras et Sakel 2007 : 842). C'est le cas de la syntaxe et notamment du changement de l'ordre des mots. Heine (2008 : 34) confirme que l'ordre des mots est particulièrement sensible au contact de langues. De plus, il est montré que dans les langues du monde, l'article défini est une catégorie qui attire la *pattern replication* plutôt que la *matter replication* (Matras et Sakel 2007 : 845). On trouve bien plus fréquemment un processus de fonctionnalisation en article défini d'un morphème hérité, existant déjà dans la langue, qu'un emprunt de l'article défini étranger tel quel.

4.2.2. Le mécanisme à l'œuvre en romani

Le romani d'Albanie a connu une suite de réplications grammaticales induites par le contact.

- L'ordre des mots canonique et hérité, DEF A N, est toujours en usage parmi les locuteurs âgés, peu instruits et parmi les Arlis (conservateurs sur le plan syntaxique).

- La polydéfinitude DEF N DEF A est un ancien ou très ancien emprunt au grec (hypothèse B), où la structure a une fonction de restriction de la dénotation du nom. Elle est utilisée par les locuteurs d'âge moyen et par les femmes. Plus exactement, le romani a effectué une convergence du type *pattern replication* en contexte contrastif. La réplication totale de l'article défini et la postposition de l'adjectif en sont les traits pivots (*pivotal features*). Cela a provoqué un changement à long terme pour le romani, qui a enrichi son inventaire structural avec le syntagme DEF N DEF A.

- L'ordre des mots DEF N A est un récent emprunt à l'albanais (hypothèse B'). Il est utilisé par les locuteurs jeunes, instruits, hommes, par les Mečkars et les Čergars I (innovateurs sur le plan syntaxique). Il s'agit d'une réplication grammaticale affectant les relations syntaxiques, où la stratégie employée est celle du « *shift from one construction type to another* ». La langue R est le romani, la langue M est l'albanais. De plus, le romani a effectué une *convergence* du type *pattern replication* de l'albanais. Le trait pivot est la simple postposition de l'adjectif (de classes 1 et 2) côté albanais, et la postposition de l'adjectif en syntagme polydéfini côté romani. L'emplacement différencié de l'adjectif en fonction de son origine confirme cette hypothèse : l'antéposition est plus fréquente pour les adjectifs hérités, la postposition plus fréquente pour les adjectif empruntés, quand bien même ceux-ci proviennent d'une langue à antéposition (cf. section 2.4.3, p. 335). La *pattern replication* de

l'albanais a été rendue possible par la *pattern replication* précédente du grec : DEF N DEF A a ainsi servi de « pont » entre DEF A N et DEF N A.

- Les locuteurs (albanais, roumains) pour qui DEF N A n'est plus un ordre marqué et qui l'emploient hors de sa fonction primaire de contraste, sont allés plus loin encore. Ils ont effectué un « démarquage » pragmatique (*pragmatic unmarking*), un processus de grammaticalisation qui fait passer du marquage pragmatique au marquage syntaxique (Heine 2008 : 54). Il consiste à sélectionner un ordre des mots pragmatiquement marqué, utilisé pour une fonction x, parce qu'il est l'ordre des mots de la langue modèle, puis à l'utiliser en vue d'une autre fonction y, qui correspond à la même fonction y dans la langue modèle (Heine 2008 : 42). La structure étend ainsi son domaine d'utilisation à une ou plusieurs autres fonctions, jusqu'à devenir non-marquée voire canonique.

4.3. Synthèse

A la lumière de nos résultats et de l'histoire du contact linguistique des différentes variétés, on constate qu'un changement de l'ordre des constituants nominaux est en cours en romani d'Albanie. On peut l'observer à l'œuvre dans la communauté arli, divisée en trois sous-groupes d'âge. La Figure 86, issue de Tirard (2015) et Tirard (2017 : 154), schématise ce scénario.

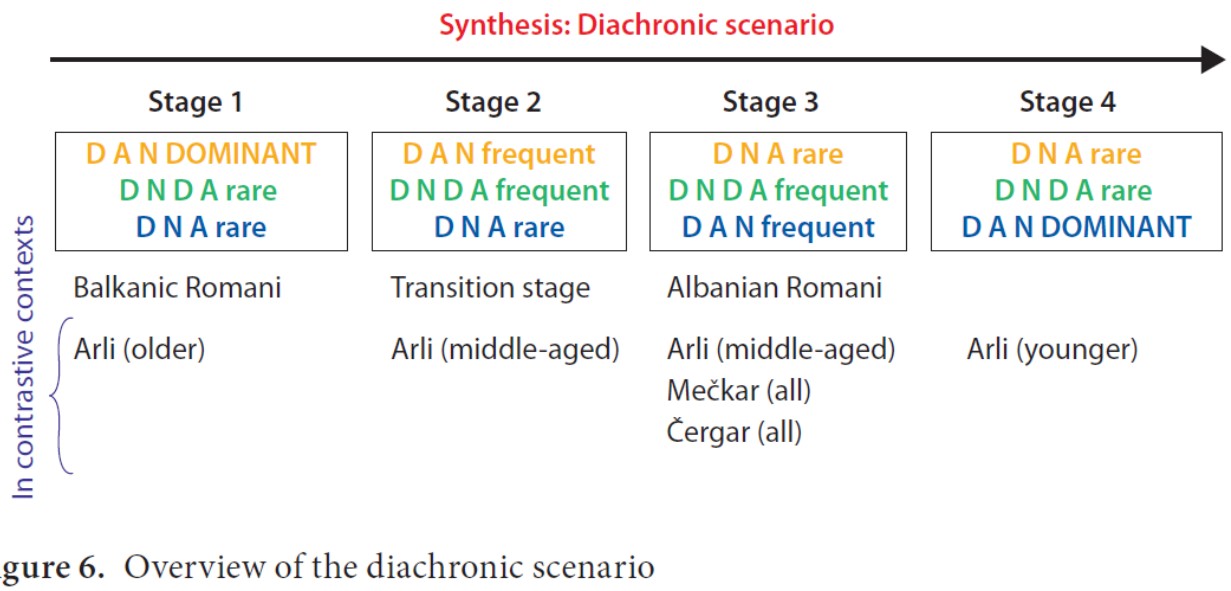


Figure 86 : Synthèse du scénario diachronique pour l'acquisition de la polydéfinitude et de la postposition de l'adjectif en romani d'Albanie (Tirard 2017 : 154)

Qu'est-ce qui a pu déclencher un tel changement ? Nous pouvons supposer que la structure canonique DEF A N a cessé d'être adéquate dans certains contextes, notamment contrastifs, et d'assurer une interprétation restrictive. Pour mieux répondre aux besoin de la communication, les locuteurs ont

donc opéré un changement interne en choisissant DEF N A. C'est notre hypothèse A' de l'innovation interne.

La structure DEF N DEF A représente un entre-deux, puisqu'il s'agit d'une structure très ancienne dans la langue (cf. sections 1.2, p. 408, et 1.5, p. 444), à l'instar de DEF A N… mais où l'adjectif est postposé, à l'instar de DEF N A. De plus, cette structure rappelle fortement celle du grec prestigieux. La structure polydéfinie DEF N DEF A a donc dû servir de « pont » pour passer de la structure initiale DEF A N à la structure nouvelle DEF N A. C'est notre hypothèse A de l'innovation interne.

Une innovation interne peut être favorisée par l'existence d'un trait proche ou similaire dans une langue de contact : le changement a été favorisé par le contact avec l'albanais dominant (qui connaît une structure « polydéfinie » très semblable, mais d'usage systématique) et avec le grec prestigieux (qui connaît une structure polydéfinie jumelle, d'usage optionnel). On a constaté que les femmes ont favorisé le changement. Ce sont nos hypothèses B et B' du contact de langues.

5. Réponse à la problématique

Nous pouvons achever, ici en brun, notre réponse à la problématique (Figure 87) :

Questions structurelles

1) Quels éléments peuvent entrer dans un syntagme polydéfini romani ?

La structure polydéfinie est constituée d'un nom précédé d'un article défini (éventuellement précédé d'une préposition) et d'un adjectif précédé d'un article défini (éventuellement précédés d'une préposition). Un déterminant peut venir remplacer l'article défini introduisant le nom : le possessif (POSS N DEF A). Le romani ne se distingue pas ici du grec.

Le corpus avec rupture interne confirme qu'il est possible d'avoir un possessif à la place du premier article défini (POSS N DEF A), mais également qu'il est possible que deux déterminants introduisent le nom, article défini et quantifieur qui lui succède (DEF QUANT N/ DEF A).

Le corpus semi-spontané montre qu'un autre déterminant peut venir précéder l'article défini en tête de syntagme : le démonstratif (DEM DEF N DEF A). Comme la polydéfinitude se construit exclusivement avec l'article défini, si l'on excepte un cas de répétition du démonstratif, mais jamais avec l'article indéfini, on peut parler d'une variable linguistique de DEFINITUDE.

2) Dans quel ordre les trouve-t-on ? Existe-t-il des structures DEF A DEF N ?

Le corpus spontané présente des syntagmes avec adjectif postposé DEF N DEF A, et quelques cas d'ordre DEF A DEF N. Le corpus avec rupture interne ne présente aucun cas de syntagme polydéfini avec adjectif antéposé.

Dans le corpus semi-spontané, l'ordre est DEF N DEF A avec adjectif postposé dans tous les cas sauf deux : un syntagme DEM DEF A DEF N et un syntagme DEF A/ DEF N avec pause interne. Le romani se distingue donc là du grec. On peut parler d'une variable linguistique d'ORDRE DES CONSTITUANTS préconisant la postposition de l'adjectif articulé.

Ayant survolé des données en romani albanais ancien, roumain ancien et d'autres régions, nous avons constaté la rareté de la structure DEF A DEF N. Elle est en perte de vitesse voire a disparu de la plupart des variétés. Nous l'expliquons par la concurrence qu'elle subit, sur le plan fonctionnel, de la part de la monodéfinitude avec adjectif postposé DEF N A.

3) L'ordre des mots est-il plus souple en syntagme poly- que monodéfini ? Le romani présente-t-il la même souplesse que le grec ?

Le romani albanais permet l'antéposition de l'adjectif (DEF A N) aussi bien que sa postposition (DEF N A) en syntagme monodéfini. Il se distingue par là du grec, qui ne permet que l'antéposition de l'adjectif (DEF A N, *DEF N A). Le romani est donc plus souple que le grec en syntagme monodéfini.

En syntagme polydéfini, DEF A DEF N comme DEF N DEF A sont possibles en romani. Toutefois, l'antéposition de l'adjectif semble marginale par rapport à sa postposition, car nous avons relevé très peu de cas d'antéposition (DEF A DEF N). Le romani se distingue par là du grec, qui antépose volontiers l'adjectif (DEF A DEF N) autant qu'il le postpose (DEF N DEF A). L'ordre des mots est donc moins souple en syntagme polydéfini que monodéfini en romani – c'est le contraire en grec.

4) Lorsqu'une expression polydéfinie contient plusieurs adjectifs, trouve-t-on les mêmes contraintes d'ordre qu'en grec ?

Le corpus spontané ne présente pas de cas de succession d'adjectifs. Tous les syntagmes polydéfinis avec rupture présentent également un seul adjectif.

Le corpus semi-spontané présente quelques cas de succession d'adjectifs. Lorsqu'une expression polydéfinie contient plusieurs adjectifs, ils peuvent encadrer le nom ou bien être tous postposés. S'ils encadrent le nom (DEF A N A et DEF A N DEF A), c'est l'adjectif postposé qui porte le contraste. Il permet de réaliser la sélection dans un ensemble de référents, plutôt que l'adjectif antéposé. Chez les locuteurs anciens et Arli, on a pu constater cependant que l'adjectif antéposé peut également réaliser la sélection d'exemples.

Si les adjectifs sont tous deux postposés, ils sont alors tous deux articulés (DEF N DEF A DEF A), ou bien seulement le premier (DEF N DEF A A), ou seulement le deuxième (DEF N A/ DEF A). Ceci contredit l'affirmation d'Igla (1996 : 166) (cf. section 2.1, p. 142), selon laquelle « *enthält die NP einen definiten Artikel, muß dieser bei Nachstellung des Adjektivs wiederholt werden* ». C'est le premier adjectif prononcé, soit le plus proche du nom, qui permet d'effectuer la sélection. Nous pouvons donc affirmer que la structure est récursive.

5) L'adjectif d'un syntagme polydéfini peut-il avoir son propre complément ?

Aucun adjectif du corpus spontané n'a son propre complément. Les syntagmes avec rupture n'en présentent pas non plus.

Aucun adjectif du corpus semi-spontané n'a son propre complément.

6) L'emplacement postnominal est-il la zone des « options » ? Quel est le degré de nominalité des éléments qui s'y trouvent ?

L'analyse du corpus semi-spontané montre que la postposition de l'adjectif, même en syntagme polydéfini, sert de position préférentielle pour les adjectif empruntés – qu'ils l'aient été à une langue antéposant ou postposant l'adjectif, qu'ils aient été morphologiquement adaptés ou non,

quelle que soit l'époque de l'emprunt. En ce sens, le romani contemporain a compartimenté sa syntaxe nominale (section 2.4.3, p. 335 et 2.4.7, p. 350).

Nous avons constaté que la postposition de l'adjectif est non-marquée dans le contexte de la tâche, où l'adjectif est restrictif. Il s'agit d'une rupture avec les autres contextes énonciatifs, voire avec les autres variétés, pour lesquelles la position non-marquée de l'adjectif est censée être antéposée.

De même, lorsque plusieurs adjectifs encadrent le nom, c'est l'adjectif postposé qui porte le contraste, ce qui confirme que la position non-marquée de l'adjectif à visée contrastive est postposée. Il convient donc de parler d'emplacements à usages différents, plutôt que d'emplacements marqués ou non-marqués.

Comme l'adjectif restrictif est nécessaire à l'identification, il n'est nullement une « option » : l'emplacement postnominal n'est donc pas une, ou pas qu'une, zone des options.

Nous pensons que le placement de l'adjectif est libre en romani albanais, et régi par des contraintes préférentielles et pragmatiques. Notre corpus albanais et les données analysées en romani de Roumanie montrent que l'adjectif postposé au nom ne porte pas toujours de marque de flexion nominale. La variation morphologique est importante sur ce plan, et remonte loin puisqu'on la trouve déjà en romani roumain du XIXème siècle.

Le romani albanais présente selon nous la même topologie syntaxique et sémantique que les langues romanes selon Cinque (2010) et Cinque (2014) : la position antéposée est celle de la modification directe, la position postposée est celle de la modification directe ou indirecte.

L'emplacement postnominal est la « post-zone » au sens de Nølke (2001). Elle peut contenir des adjectifs épithètes, des noms en apposition étroite, des syntagmes génitivaux et même des déterminants. Ces derniers (article défini, démonstratif, possessif, numéral) peuvent accompagner un adjectif, un nom en apposition étroite… ou rien du tout lorsqu'ils clôturent le syntagme (structures DEF N DEF A DEM, QUANT DEF N DEM). Cela appelle une réflexion sur le nom : la présence d'un article défini à côté d'un élément n'en fait pas nécessairement un nom. Les cas de détermination multiple relevés dans les variétés albanaise et roumaine du romani montrent qu'un nom n'est pas toujours la tête d'un seul déterminant.

La post-zone permet un emploi spécifique, discriminant, classificateur de l'adjectif. Elle est le lieu préférentiel d'emploi des adjectifs empruntés – et dont les locuteurs ont conscience du caractère étranger. Elle permet un focus étroit sur l'adjectif postposé, et probablement aussi sur le syntagme génitival postposé. Le focus interne au syntagme nominal réside dans un emplacement syntaxique donné. Ainsi proposons-nous un schéma « géographique » du syntagme nominal en romani d'Albanie, en Figure 83.

Figure 83 : Topologie du syntagme nominal en romani d'Albanie

7) Existe-t-il une restriction quant au type de nom pouvant entrer dans un syntagme polydéfini ?

Nous avons trouvé dans le corpus spontané presque tous les types de noms : un nom propre et des noms communs, un nom indénombrable et des noms dénombrables, un nom d'événement, des noms abstraits, animés, inanimés. Nous n'avons pas trouvé de nom massif ni de nom générique. Ils ont des poids différents et sont aussi bien hérités qu'empruntés.

Les syntagmes avec rupture ne présentent que des noms communs, dénombrables et animés spécifiques. Ils sont de poids léger et d'origine héritée.

Nous n'avons pas trouvé de contraintes quant au type de nom dans le corpus semi-spontané.

8) Existe-t-il des contraintes morphologiques ou étymologiques quant au type de l'adjectif ?

Le corpus spontané ne présente que des adjectifs hérités. Nous n'avons pu établir de corrélation entre leur poids et leur position dans le syntagme. Les adjectifs qualificatifs des syntagmes avec rupture sont tous hérités.

En corpus semi-spontané, nous n'avons pas trouvé de contrainte de morphologie, de poids ni d'origine étymologique pour l'adjectif.

9) Existe-t-il des contraintes syntaxiques ? L'adjectif doit-il être prédicatif ?

Le corpus spontané ne présente que des adjectifs prédicatifs, à l'exception de *vaver*, autre. Les adjectifs qualificatifs des syntagmes avec rupture sont tous prédicatifs.

La nature du corpus semi-spontané fait que tous les adjectifs sont prédicatifs (à l'exception de « autre »).

10) Existe-t-il des contraintes sémantiques ? Doit-il être restrictif ? Certains adjectifs sont-ils exclus, comme les adjectifs relationnels ou ceux lexicalisés et figés avec le nom ?

Les adjectifs articulés sont orientés vers l'objet et non vers le locuteur, restrictifs et intersectifs (ou intersectifs en apparence subsectifs). Ils permettent une interprétation conjointe. Nous avons toutefois observé plusieurs cas où le corpus polydéfini présente un adjectif non-restrictif, notamment lorsqu'il sert à caractériser, à mettre une étiquette (*epithet*). Dans d'autres cas, on peut parler de restriction seulement si l'on considère un ensemble plus large que celui des seuls référents possibles du nom.

Le corpus spontané présente un cas d'adjectif relationnel et deux cas d'adjectifs semi-lexicalisés avec le nom, ce qui montre qu'ils ne sont pas exclus de la polydéfinitude. Toutefois, dans les deux cas, ils impliquaient un contraste implicite ou explicite.

Les syntagmes avec rupture présentent des adjectifs qualificatifs permettant une interprétation intersective conjointe, tous orientés vers l'objet. Plusieurs sont restrictifs, mais d'autres sont

ambigus quant à une interprétation restrictive ou non : un cas de l'adjectif de caractérisation et deux cas de restriction dans un ensemble plus large que les seuls référents possibles du nom.

La nature du corpus semi-spontané fait que tous les adjectifs articulés y sont restrictifs. Ils ont tous une fonction « modificatrice » dans le sens de Nølke (2001) et donnent lieu à une interprétation intersective du syntagme nominal. L'analyse nous a également permis de déterminer la fonction des différentes structures syntagmatiques dans le contexte de la tâche. DEF N DEF A est la structure privilégiée à part égale par les adjectifs de toutes origines. Elle est en revanche la structure privilégiée des adjectifs à lecture intersective en apparence subsective, qui se trouve être hérités (taille, *autre*). Pour comprendre la polydéfinitude, on voit qu'il faut donc combiner une explication par l'origine, la position et la sémantique de l'adjectif. C'est la variable du TYPE D'ADJECTIF.

11) Existe-t-il des contraintes quant au type de déterminant ? Peut-il y avoir répétition d'autre chose que de l'article défini, notamment l'article indéfini ?

Seul l'article défini est répliqué dans notre corpus ; aucun déterminant indéfini ne semble pouvoir être répété. Le possessif peut prendre la place de l'une des occurrences de l'article défini – mais il n'est pas pour autant répliqué. C'est l'article défini qui est alors répété : POSS N DEF A. La structure ne se produit donc qu'en contexte défini : un déterminant défini simple ou complexe en première position et un article défini en seconde position.

Les syntagmes à rupture interne nous amènent à nuancer cette dernière affirmation, puisque l'on peut trouver en première position un déterminant défini associé à un déterminant indéfini, le quantifieur : structure DEF QUANT N/ DEF A.

Seul l'article défini est répliqué dans le corpus semi-spontané, à l'exception d'une occurrence de démonstratif (DEM N DEM A).

Nous avons relevé une occurrence de réduplication du démonstratif dans un texte en romani roumain du XIXème siècle (DEM N DEM A), une occurrence de réduplication du possessif dans un corpus contemporain de Serbie (POSS N POSS A) et deux occurrences de réplication du possessif dans une étude portant sur le romani contemporain de Grèce (POSS A POSS N et POSS NUM POSS N où le numéral est analysé comme un adjectif). L'étude des autres variétés de romani montre que ce sont celles de Roumanie qui sont les plus créatives en matière de surdétermination.

- Le premier déterminant du syntagme peut être un démonstratif (DEM N DEF A), un possessif (POSS N DEF A), un quantifieur (QUANT N DEF A), tandis que le deuxième déterminant doit être un article défini… à l'exception des structures roumaines DEF N DEM A et POSS N DEM A, où l'adjectif n'est toutefois pas accompagné d'un article défini.
- Nous avons relevé des cas de triple détermination, avec réduplication de l'article défini associé à un démonstratif (DEF N DEM DEF A et sa variante DEF N DEF A DEM) ou à un possessif (DEF N POSS DEF A). On trouve également la triple détermination avec article défini, possessif et numéral (POSS N DEF NUM A et sa variante POSS NUM N DEF A).
- Nous avons relevé des cas de quadruple détermination avec réduplication de l'article défini associé au possessif et au numéral (DEF NUM DEF A POSS N)… et même avec réduplication du possessif associé à un article défini et à un numéral (POSS NUM POSS N DEF A).

Il n'y a donc pas à proprement parler de *poly-définitude*, en ce sens qu'il n'y a pas nécessairement réplication de l'article défini donc plusieurs instances d'article défini, dans les cas où le premier déterminant est un possessif ou un démonstratif (POSS N DEF A, DEM N DEF A). Certes, il y a plusieurs instances de déterminant défini, simple ou complexe, et le déterminant qui « accompagne » l'adjectif est l'article défini.

12) L'article défini du romani est-il explétif ?

Le corpus spontané montre que l'article défini assume toutes les fonctions présentées par Boretzky (2000b : 44-54). Dans les syntagmes polydéfinis à adjectif antéposé, il revêt un usage sémantique (*knowledge-dependent*) ; dans ceux à adjectif postposé, il revêt aussi bien un usage sémantique que pragmatique (*situation-dependent*).

L'article défini supplémentaire, qui porte sur l'adjectif, joue le rôle d'un opérateur d'identification ou de restriction. Nous considérons qu'il ne s'agit pas d'un « véritable » déterminant, mais d'un morphème de focalisation. L'article défini adjectival n'a pas les caractéristiques habituelles du déterminant, car il ne satisfait pas les contraintes d'unicité, d'existence ni de familiarité. Nous choisissons donc de dissocier article défini et définitude, comme le proposent Lekakou et Szendrői (2012 : 116) pour le grec. En ce sens, on peut parler de « marquage multiple » de la définitude plutôt que de définitude multiple ou de polydéfinitude.

13) Peut-il y avoir répétition du marquage de cas et de la préposition ?

La polydéfinitude donne lieu à la répétition de la préposition lorsqu'elle a fusionné avec l'article défini. Si un syntagme polydéfini est introduit par une préposition monosyllabique cliticisée à DEF_1, on la trouve cliticisée à DEF_2, donnant lieu à une structure PREP.DEF_1 N PREP.DEF_2 A.

Le corpus avec rupture interne montre un cas de réplication des marques flexionnelles : l'un des exemples présente une marque de flexion sur le nom et son déterminant, qui est répétée sur l'adjectif et son déterminant postposés. On en déduit que lorsque le syntagme porte une marque de cas, celle-ci est nécessairement répétée, du moins lorsque l'adjectif est postposé.

Le corpus semi-spontané comprend lui aussi plusieurs cas de répétition des marques flexionnelles et une répétition de la préposition.

14) La partie adjectivale est-elle un cas de modification directe ou indirecte ?

Nous avons analysé la partie adjectivale des syntagmes polydéfinis comme un cas de modification indirecte.

La différence entre modification directe et indirecte est d'ailleurs apparue également dans les syntagmes avec rupture et nous avons analysé la partie adjectivale des syntagmes polydéfinis comme de la modification indirecte.

15) Elle-t-elle un cas d'apposition étroite ?

Si l'on conçoit la différence entre apposition étroite et apposition lâche comme une différence structurelle et non sémantique, alors il existe la même distinction en romani. L'apposition étroite, dont ressortent les syntagmes polydéfinis sans rupture, constitue une seule unité prosodique, portant un seul accent. L'apposition lâche, dont ressortent les syntagmes polydéfinis avec rupture, comprend deux sous-unités prosodiques et porte deux accents nucléaires. Le syntagme polydéfini est une apposition étroite, comprenant une partie nominale et une partie adjectivale.

En revanche, il existe une différence importante entre les syntagmes polydéfinis lâches du romani et ceux du grec moderne. En romani ils peuvent être restrictifs (à l'instar des polydéfinis étroits) alors que ceux du grec ne sont pas restrictifs, si l'on en croit la littérature. Il ne s'agit donc pas, en réalité, du même type d'apposition. Plusieurs de nos exemples constituent un autre type encore que celui des appositions lâches de la littérature : restrictifs comme les polydéfinis classiques (apposition étroite), ils comprennent deux sous-unités prosodiques distinctes comme les polydéfinis lâches (apposition lâche). Pour autant, il ne s'agit pas de deux syntagmes juxtaposés DEF N DEF N où l'adjectif est substantivé. En effet, lorsque l'adjectif substantivé est introduit par un article défini, il n'est alors pas restrictif. Lorsque l'adjectif substantivé est restrictif, il n'est alors pas introduit par un article défini, mais par un démonstratif.

16) Existe-t-il des syntagmes polydéfinis lâches ? Leur interprétation est-elle non-restrictive ?

Il existe des syntagmes polydéfinis lâches en romani comme en grec. Ils présentent une rupture entre la sous-partie « nominale » et la sous-partie « adjectivale » : en général une pause, avec parfois un ensemble de mots insérés. Il peut s'agir d'un ou de plusieurs mots : nom propre, interjection, proposition relative, voire le reste de l'énoncé. Cette diversité rappelle celle du grec. L'un de nos exemples montre que le syntagme peut également ne pas présenter de pause.

Dans notre corpus, l'interprétation des syntagmes polydéfinis avec rupture interne est moins systématiquement restrictive que celle des syntagmes polydéfinis sans rupture interne. Il y a donc une différence avec le grec, où un syntagme avec pause ou insertion implique que l'adjectif apporte une information non essentielle et non-restrictive.

17) Y a-t-il un lien entre polydéfinitude et ordre des constituants nominaux ? Lequel ?

Le recensement des syntagmes polydéfinis en romani d'Albanie, d'autres régions ou d'autres époques a montré qu'il existe un lien avec l'ordre des constituants nominaux. Contrairement au grec, la polydéfinitude du romani se manifeste principalement lorsque l'adjectif est postposé. Or l'analyse a montré également qu'en position postposée, l'adjectif peut recevoir un focus étroit (contrairement à la position antéposée), et que l'article défini revêt des usages différents en syntagme polydéfini à adjectif antéposé ou postposé.

Toutes les variétés qui comportent des syntagmes polydéfinis peuvent postposer l'adjectif en syntagme monodéfini, de façon marquée ou non, dans un certain but communicatif ou dans certaines tournures. C'est donc que la postposition de l'adjectif existe dans la grammaire de la langue. Or on sait que les relations syntaxiques font facilement l'objet d'un transfert linguistique.

Nous pensons que la postposition de l'adjectif a été favorisée par le contact de langues, notamment pour les variétés en contact avec des langues postposant l'adjectif de façon canonique.

Notre étude contredit donc l'affirmation de Matras (2000b : 103-104), selon laquelle la postposition de l'adjectif postposé rend obligatoire la présence d'un article défini l'accompagnant (DEF N DEF A et DEM N DEF A, mais *DEF N A et *DEM N A).

Questions sémantiques

1) Existe-t-il une nuance de sens entre syntagme monodéfini et syntagme polydéfini ?

Les syntagmes polydéfinis donnent l'impression d'être plus spécifiques, probablement parce qu'ils sont restrictifs. Les exemples de syntagmes polydéfinis avec rupture fournis par notre corpus sont en revanche tout aussi spécifiques que les syntagmes monodéfinis avec rupture.

2) *Poly*définitude implique-t-elle *sur*définitude ?

La polydéfinitude n'implique pas, à notre sens, de sur-définitude. On pourrait aller jusqu'à dissocier définitude et article défini, en qualifiant celui-ci d'explétif. Il pourrait s'agir d'un marqueur de focalisation. Les syntagmes avec rupture ne se différencient pas des syntagmes sans rupture.

Il n'y a pas à proprement parler de *polydéfinitude*, en ce sens que le syntagme n'est pas « plus » défini sur le plan sémantique : le deuxième article défini n'a pas pour fonction de rendre plus défini le syntagme nominal.

3) Les syntagmes monodéfinis sont-ils restrictifs, tandis que les syntagmes polydéfinis peuvent être restrictifs ou non-restrictifs ?

En syntagme monodéfini comme en syntagme polydéfini, l'interprétation est soit restrictive soit non-restrictive. Le syntagme polydéfini a pour caractéristique d'effectuer une restriction pragmatique parmi les référents possibles du nom. On peut parler d'une fonction d'identification, qui est la fonction majeure de cette structure. Elle est donc notamment employée lorsque les référents possibles du nom sont plus de deux, et que l'identification est alors considérée comme délicate. Dans les cas (beaucoup moins nombreux) où le syntagme a un effet d'emphase, cette fonction d'identification due à la restriction pragmatique ne semble pas être présente.

Les syntagmes avec rupture ne se différencient pas des syntagmes sans rupture : l'interprétation est soit restrictive soit non-restrictive, que le syntagme soit monodéfini ou polydéfini.

4) Si l'adjectif est antéposé au nom, l'interprétation est-elle seulement restrictive ? Si l'adjectif est postposé au nom, l'interprétation peut-elle être restrictive ou non-restrictive ?

En syntagme polydéfini, lorsque l'adjectif est antéposé, l'interprétation est plutôt non-restrictive ; lorsque l'adjectif est postposé, elle est plutôt restrictive. Mais ce ne semble pas être une règle absolue. En syntagme monodéfini, lorsque l'adjectif est antéposé, l'interprétation est restrictive ou non-restrictive ; lorsque l'adjectif est postposé, elle est plutôt restrictive. Là aussi, il ne s'agit que de tendances. C'est donc la position de l'adjectif plutôt que la réplication de l'article défini qui semble jouer un rôle.

Les syntagmes polydéfinis avec rupture présentent tous un adjectif postposé, et une interprétation soit restrictive, soit non-restrictive. Il est donc difficile d'en déduire une tendance. En syntagme monodéfini, lorsque l'adjectif est antéposé, l'interprétation non-restrictive ; lorsque l'adjectif est postposé, elle est restrictive. La taille du corpus est très restreinte, mais elle confirme l'importance de l'ordre des constituants.

Questions fonctionnelles

1) La polydéfinitude est-elle optionnelle en romani comme en grec ? Si oui, qu'est-ce qui provoque le choix du locuteur ?

Le syntagme polydéfini est optionnel en romani comme en grec.

Les variables linguistiques pertinentes sont le contexte défini (DEFINITUDE), la présence d'un contraste implicite ou explicite et la postposition de l'adjectif (ORDRE DES CONSTITUANTS), l'emploi restrictif ou de caractérisation de l'adjectif (TYPE D'ADJECTIF).

La structure du syntagme nominal en romani ne présente pas que des contraintes catégoriques, mais également des contraintes préférentielles (telles que définies par Thuilier 2015) : la polydéfinitude et la postposition de l'adjectif en syntagme monodéfini font l'objet de ces dernières. Notre étude contredit donc l'hypothèse selon laquelle la présence de l'article défini après le nom s'explique uniquement par la présence de l'adjectif après le nom, qui ne peut pas occuper seul cette position (Matras 2000b : 103-104). L'ajout de l'article défini « adjectival » fait l'objet d'un choix. La postposition de l'adjectif (DEF N A et DEF N DEF A) permet de mettre l'adjectif sous focus étroit, ce que ne permet pas l'antéposition. L'adjonction d'un article défini supplémentaire constitue une mise sous focus de l'adjectif. Elle augmente le statut informatif de celui-ci et contribue ainsi à l'identification du référent. L'usage de la polydéfinitude marque la prise en compte de la relation interlocutive. Il montre que le locuteur se préoccupe de la manière dont son interlocuteur pensera le référent. L'usage de la monodéfinitude révèle au contraire que le locuteur n'a pas conscience de la présence d'éventuels référents « rivaux » dans la dénotation du nom.

2) S'agit-il d'une variation optionnelle de type *syntactic doubling* (Barbiers 2013 : 7) ?

Ce phénomène est très proche du redoublement syntaxique (*syntactic doubling*) mais cela n'en est pas un, car la redondance de l'article défini contribue à l'interprétation pragmatique du syntagme.

3) Quelle est la fonction des éléments situés dans l'emplacement postnominal (épithète, apposition, *afterthought*…) ?

Le syntagme polydéfini a pour fonction l'identification. C'est pourquoi il comprend, le plus souvent, un adjectif postposé : la postposition est le lieu du focus, où les sèmes spécifiques sont mis en avant. Nous avons vu qu'un syntagme polydéfini peut avoir d'autres fonctions : précision ou auto-correction, récapitulation, *topic-shift*.

Les syntagmes avec rupture ne se différencient pas des syntagmes sans rupture.

Les éléments situés dans l'emplacement postnominal peuvent être des modifieurs épithètes focalisés, selon le résultat de notre étude, ainsi que des appositions lâches. Cela n'empêche pas d'y trouver également des « options » ayant la fonction d'*afterthought* (Matras 2002 : 166).

L'emplacement postnominal est une zone qui contient des éléments autres que les seules options et les *afterthoughts*. On y trouve des appositions étroites (nominales ou adjectivales, qui font les syntagmes polydéfinis), des appositions lâches (adjectivales ou nominales), des syntagmes génitivaux. En ce qui concerne les syntagmes polydéfinis, nous avons pu relever des fonctions de contraste ou de sélection, d'emphase, de précision sur un référent mentionné, de synthèse du discours qui précède, de surnom et de romanisation d'un nom, d'un terme ou d'un concept étranger.

4) Le syntagme polydéfini implique-t-il un contraste ?

Le focus implique souvent un contraste, mais celui-ci n'est pas nécessairement explicite. En effet, on a pu constater que l'effet de sélection de DEF N DEF A va de pair avec un contraste *in presentia* (référents du même ensemble mentionnés en discours ou présents dans la situation d'énonciation) ; l'effet d'emphase de DEF N DEF A va de pair avec un contraste *in absentia*.

On remarque que le contraste *in presentia*, qui correspond à l'effet de sélection explicite, se trouve plus dans les syntagmes avec rupture que sans rupture.

L'ensemble des syntagmes polydéfinis du corpus spontané implique un contraste, car la tâche à réaliser a été conçue dans ce dessein.

5) Quel est le statut de l'information qu'il porte ?

Le statut de l'information du nom n'est pas pertinent, dans la mesure où les syntagmes polydéfinis ne transmettent pas nécessairement une information nouvelle, mais aussi une information donnée ou accessible. L'adjectif, lui, est toujours nouveau lexicalement : il apporte donc une information nouvelle, une propriété qui permet l'identification du référent. Les syntagmes avec rupture ne se différencient pas des syntagmes sans rupture.

6) Fait-il plutôt partie du topique ou du focus de la phrase ?

Les syntagmes polydéfinis peuvent faire partie soit du topique, soit du focus de la phrase. Les syntagmes avec rupture ne se différencient pas des syntagmes sans rupture.

7) Implique-t-il un focus sur l'adjectif ? Si oui, de quel type ?

Le syntagme polydéfini implique un focus sur la partie adjectivale du syntagme, soit informatif soit contrastif – sans que l'on ait pu déceler de lien avec la position de l'adjectif. Les syntagmes avec rupture ne se différencient pas des syntagmes sans rupture.

La présence d'un article défini devant l'adjectif augmente le statut informatif de celui-ci, de même que son emplacement après le nom. L'adjectif antéposé au nom ne porte pas de focus propre, tandis que c'est le cas de l'adjectif postposé. Il existe des langues où les mécanismes de focalisation combinent déplacement du terme focalisé et adjonction d'un marqueur de focalisation (Creissels 2006 : 126) : le romani fait partie de ces langues. Nous pensons que la focalisation n'est pas un phénomène binaire mais scalaire, et que les syntagmes monodéfinis, polydéfinis, à adjectif antéposé ou postposé, se répartissent le long de cette échelle.

8) Le focus est-il situé sur le la totalité du syntagme polydéfini ou sur sa partie adjectivale ? Dans le deuxième cas, la partie nominale constitue-t-elle la présupposition ?

Il apparaît que c'est la topologie du syntagme nominal qui détermine quelle sous-partie constitue la présupposition, et quelle partie constitue le focus. La place postnominale est une position de focus, ou rhématique, tandis que la position anténominale est une position non-focale, ou thématique. La première est le lieu de la modification indirecte, tandis que la seconde est le lieu de la modification directe. Du moins est-ce le cas dans les variétés mečkar et čergar I, la variété arli présentant (encore) une distinction moins nette entre les deux. On peut ainsi représenter les quatre structurations possibles d'un syntagme nominal avec adjectif épithète sans rupture interne (Tableau 40). Les syntagmes avec rupture ne se différencient pas des syntagmes sans rupture.

Tableau 40 : Quatre structurations possibles du syntagme nominal avec adjectif épithète

	Adjectif antéposé	Adjectif postposé
Sans répétition ART.DEF	DEF A N *o baro Del, to purane breša* adjectif en rôle thématique (présupposé) article défini restrictif ou non	DEF N A *o pete bare, o djalekto romano* adjectif en rôle rhématique (focus) article défini restrictif ou non
Avec répétition ART.DEF	DEF A DEF N *o tikne o maksemi, o barvalo o Skender* adjectif en rôle thématique (présupposé) article défini restrictif	DEF N DEF A *mo phral o tikno, o filme o bare* adjectif en rôle rhématique (focus) article défini restrictif

L'article défini « adjectival » permet de focaliser l'adjectif. Il joue le rôle d'un opérateur d'identification en renforçant le statut référentiel de l'ensemble du syntagme. On peut dire qu'il joue le rôle d'un morphème de focalisation car il permet le focus sur la partie adjectivale du syntagme nominal. L'adjectif est souvent nouveau sur le plan du statut de l'information ; la partie nominale peut avoir un statut donné ou accessible, parfois nouveau.

Questions socio-linguistiques

1) Tous les locuteurs albanais du romani emploient-ils la structure polydéfinie ?

Tous les locuteurs albanais du corpus semi-spontané emploient la structure polydéfinie… à l'exception d'une personne ayant le profil socio-linguistique suivant : homme, non-militant, âgé de plus de 60 ans, parlant la variété arli (dans sa famille d'origine comme dans sa famille actuelle avec son épouse et ses enfants), ayant un niveau d'instruction très faible (école primaire seulement) et bilingue en grec. Ce locuteur possède toutes les caractéristiques pour antéposer l'adjectif et pratiquer peu la polydéfinitude.

Mais étonnamment, ce locuteur a employé la structure polydéfinie dans d'autres occasions que lors de la tâche semi-spontanée. Notre corpus spontané en recèle deux occurrences : c'est lui qui a prononcé les exemples (158) et (170). On peut donc en déduire que la totalité des locuteurs albanais que nous avons rencontrés et enregistrés emploie la polydéfinitude.

2) Observe-t-on une différence de profil socio-linguistique parmi les locuteurs employant cette structure : différence de genre, d'âge ou de niveau d'éducation ?

Le corpus spontané permet quelques premières constatations quant au profil des locuteurs pratiquant la polydéfinitude. Il s'agit aussi bien d'hommes que de femmes, appartenant à toutes les classes d'âge (sauf les plus jeunes ?).

Le corpus avec rupture confirme que la polydéfinitude est pratiquée par tous les locuteurs du romani d'Albanie : homme et femmes de toutes classes d'âge.

L'analyse du corpus semi-spontané a montré que les variables MILITANTISME et EDUCATION sont moins pertinentes que les autres pour expliquer l'usage de la polydéfinitude. En ce qui concerne le GENRE, on constate que les femmes emploient plus la polydéfinitude que les hommes. Les deux variables les plus importantes sont AGE et VARIETE. Le corpus semi-spontané a confirmé les résultats du corpus spontané, à savoir que la polydéfinitude est peu pratiquée par les plus jeunes. Elle est le plus pratiqué par les adultes ayant entre 15 et 59 ans.

3) Quelles variétés présentent le plus cette structure ?

Dans le corpus spontané, la polydéfinitude est employée par des locuteurs de plusieurs variétés : arli, čergar I, mečkar, arli & čergar I.

Dans le corpus spontané avec rupture interne, la polydéfinitude est employée par des locuteurs de

plusieurs variétés : arli, mečkar, čergar I & arli, mečkar & čergar I.

Le corpus semi-spontané confirme qu'il existe une variation dialectale quant à l'usage de la polydéfinitude. Les variétés qui l'emploient le plus sont le mečkar et le čergar I. Celles qui postposent le plus l'adjectif, en syntagme monodéfini ou polydéfini, sont le mečkar et le mečkar bi-variétal.

La variable de la VARIETE doit être corrélée à celle de l'AGE. On distingue les groupes suivants :

- Les Mečkars et Čergars I connaissent la *stabilité* linguistique, car ils ont achevé le changement de l'ordre initial DEF A N vers DEF N DEF A puis l'ordre innovateur DEF N A.
- Les Arlis âgés, conservateurs sur le plan syntaxique, connaissent également la *stabilité* car ils n'ont pas opéré de changement.
- Les Arlis d'âge moyen expérimentent le modèle du *changement au cours de la vie* car ils ont adopté, au cours de leur vie, le changement en question, mais sans l'avoir systématisé.
- Les jeunes Arlis emploient massivement la variante innovante, qu'ils correspondent au modèle de la *gradation par l'âge* ou à celui de *changement générationnel*.

4) Les locuteurs qui emploient cette structure peuvent-ils être influencés par la connaissance d'une ou plusieurs langues étrangères, et si oui, lesquelles ?

Tous les locuteurs de notre étude ont comme langue seconde l'albanais. Ceux qui emploient la polydéfinitude peuvent ne connaître aucune autre langue, ou bien connaître également : le grec, l'italien, le serbe, le macédonien ou le bulgare. D'autres ont des connaissances scolaires, télévisuelles ou militantes de l'anglais, de italien et de l'espagnol.

Le grec est une langue à polydéfinitude qui peut fortement influencer les locuteurs de notre corpus. Toutefois, c'est également une langue qui antépose l'adjectif et prohibe la postposition de l'adjectif (non-articulé), ce qui peut avoir l'influence inverse. L'italien et l'espagnol ne présentent pas de polydéfinitude mais sont des langues où la modification indirecte correspond à la postposition de l'adjectif : elles ont pu jouer un rôle.

En anglais au contraire, la modification indirecte correspond à l'antéposition de l'adjectif. De même, le serbe, le macédonien et le bulgare sont des langues slaves antéposant l'adjectif épithète. La connaissance de ces langues n'a pu favoriser l'emploi de la polydéfinitude. Pour en avoir le cœur net, on peut également effectuer une comparaison avec les témoins, locuteurs natifs du macédonien et du serbe. Leur syntaxe nominale est bel et bien fort différente :

- la polydéfinitude est quasiment inexistante (DEFINITUDE).
- la postposition de l'adjectif est quasiment inexistante (ORDRE DES CONSTITUANTS).
- l'antéposition de l'adjectif est non-marquée dans le contexte contrastif de la tâche.
- lorsque plusieurs adjectifs encadrent le nom, c'est l'adjectif antéposé qui porte le contraste, ce qui confirme que la position non-marquée de l'adjectif à visée contrastive est antéposée.

5) Quel est le profil « type » de la personne employant le plus volontiers la structure polydéfinie ?

La personne qui a le plus de probabilité de produire fréquemment DEF N DEF A est une femme, non-militante, âgée entre 15 et 59 ans et parlant la variété mečkar ou čergar I.

6) Quel est le profil « type » de la personne postposant le plus volontiers l'adjectif ?

Dans le corpus spontané, la monodéfinitude avec adjectif postposé est pratiquée par une plus grande diversité de locuteurs, d'âges et de variétés différentes.

Le corpus semi-spontané confirme qu'il existe une variation dialectale quant au placement de l'adjectif épithète en Albanie. L'analyse montre que la postposition de l'adjectif en syntagme monodéfini est le fait des locuteurs les plus jeunes, plutôt locuteurs des variétés mečkar ou čergar I. Plus exactement, la personne qui a le plus de probabilité de produire fréquemment DEF N A est un homme de moins de 15 ans, en cours de scolarisation (ou moyennement scolarisé s'il est plus âgé), non-militant, parlant čergar I ou čergar I bi-variétal (avec mečkar ou arli), ou encore mečkar. La personne qui a le plus de probabilité de postposer l'adjectif (DEF N A ou DEF N DEF A) est un homme ou une femme de moins de 15 ans, non-militant, parlant mečkar ou mečkar bi-variétal (avec čergar I ou bamlija).

Le placement des adjectifs sur l'échelle d'héritage (ou *inheritedness*, « héritéité ») en fonction de leur origine laisse à penser que la structure la plus innovative est DEF N A, et que DEF N DEF A constitue elle aussi une innovation, pour ainsi dire plus ancienne.

7) Quel est le profil « type » de la personne antéposant le plus volontiers l'adjectif ?

Dans le corpus spontané, la monodéfinitude avec adjectif antéposé est pratiquée quasiment exclusivement par les Arlis et les locuteurs de plus de 40 ans.

Dans le corpus avec rupture interne, la monodéfinitude avec adjectif antéposé est le fait d'un profil beaucoup plus restreint de locuteurs que celle avec adjectif postposé : locuteurs arli d'âge moyen ou avancé dans le premier cas, locuteurs de toutes variétés et de tous âges dans le deuxième cas.

L'analyse du corpus semi-spontané confirme que l'antéposition de l'adjectif, mono- ou polydéfini, est le fait des locuteurs de plus de 40 ans et parlant arli. Plus exactement, la personne qui a le plus de probabilité de produire fréquemment DEF A N a plus de 60 ans, est un homme ou une femme militant, qui parle arli, arli bi-variétal ou est un néo-locuteur arli. La personne qui a le plus de probabilité d'antéposer l'adjectif (DEF A N ou DEF A DEF N) est un homme ou une femme de plus de 40 ans, militante, qui parle arli, arli bi-variétal ou est un néo-locuteur arli.

Questions diachroniques

1) À quand remonte cette structure ? Est-il possible de la dater ?

Des syntagmes polydéfinis sont attestés dans des textes en romani albanais des années 1930 et dans des textes en romani roumain du milieu du XIXème siècle. Nous pensons que cette structure est très ancienne, simultanée ou postérieure à l'acquisition de l'article défini, qui est un balkanisme majeur du romani. Elle remonte probablement au vieux romani, donc à l'arrivée des Roms en Europe (XIème siècle ?), auquel cas toutes les variétés actuelles ont un jour connu la polydéfinitude dans leur histoire. Celles parlées dans des zones où les langues de contact ne présentent pas d'article défini ont commencé par perdre la polydéfinitude, avant de perdre progressivement l'article défini lui-même. Celles parlées dans des zones où les langues de contact présentent un article défini mais pas de polydéfinitude ont perdu l'usage de la polydéfinitude mais ont conservé l'article défini. En l'absence d'un véritable corpus historique, on ne peut toutefois exclure que la polydéfinitude soit postérieure et se limite à la seule région des Balkans et du centre-est de l'Europe.

Le romani des Balkans pratique la postposition de l'adjectif depuis une époque très ancienne, postérieure à celle du vieux romani mais antérieure au XIXème siècle. En romani d'Albanie, l'acquisition de la postposition de l'adjectif (en syntagme monodéfini, polydéfini ou indéfini) est postérieure à l'acquisition de la polydéfinitude (à adjectif antéposé ou postposé).

2) D'où provient-elle ? S'agit-il d'une innovation interne ou d'un « emprunt » ?

La polydéfinitude du romani provient d'une langue de contact, probablement du grec, langue de grande influence sur le vieux romani et le romani médiéval balkanique, avec lequel sa polydéfinitude présente le plus de similarités formelles et fonctionnelles. D'autres langues locales présentant une forme de polydéfinitude, comme le roumain ou l'albanais, ont pu favoriser son maintien ou son développement de cette structure.

3) Quel a été le processus d'acquisition de la structure ?

L'acquisition des syntagmes polydéfinis est un cas de transfert linguistique par contact de langues, ainsi que de convergence linguistique de type *pattern replication*, en contexte restrictif et contrastif. Les traits pivots sont la réduplication totale de l'article défini et la postposition de l'adjectif. La langue R est le romani, la langue M est le grec. Cette réplication grammaticale a provoqué un changement à long terme pour le romani, en enrichissant son inventaire structural des syntagmes DEF N DEF A et DEF A DEF N.

L'acquisition de la postposition de l'adjectif en syntagme monodéfini est un transfert linguistique, de type réplication grammaticale affectant les relations syntaxiques, où la stratégie employée est celle du *shift from one construction type to another*. Il s'agit également d'une convergence linguistique, de type *pattern replication* où le trait pivot est la postposition de l'adjectif (de classes 1 et 2) côté albanais, et la postposition de l'adjectif en syntagme polydéfini côté romani. La langue R est le romani, la langue M est l'albanais, le roumain, ou toute autre langue de contact dominante

locale qui postpose l'adjectif. Cette *pattern replication* de l'albanais a été rendue possible par la *pattern replication* précédente du grec, c'est-à-dire que DEF N DEF A a permis l'acquisition de DEF N A, en servant d'intermédiaire entre DEF A N (structure héritée) et DEF N A. En effet, la structure DEF N DEF A représente un entre-deux : à l'instar de DEF A N, il s'agit d'une structure très ancienne dans la langue, et à l'instar de DEF N A, l'adjectif y est postposé. De plus, cette structure ressemble à celle que l'on trouve en grec (prestigieux).

La généralisation de la postposition de l'adjectif en syntagme monodéfini par les locuteurs de certaines variétés est un transfert linguistique, de type réplication grammaticale affectant les relations syntaxiques, où la stratégie employée est celle du *pragmatic unmarking*.

S'agissant des différentes variétés de romani balkanique, le processus a dû être le suivant. Le romani balkanique, qui avait déjà DEF A N (hérité), a emprunté à la fois DEF A DEF N et DEF N DEF A au grec. Il a ensuite progressivement perdu DEF A DEF N, dans les zones où la langue de contact ne le pratique pas (Roumanie, Albanie). Par ailleurs il a effectué un deuxième emprunt, DEF N A (aux langues roumaines, qui pratiquent à la fois DEF N A et DEF N DEF A).

Là où l'emprunt de DEF N A était rare (Albanie), il s'est appuyé sur la langue de contact pour créer une innovation interne. Nous pouvons supposer que la structure canonique DEF A N a cessé d'être adéquate dans certains contextes et d'assurer une interprétation restrictive. Pour mieux répondre aux besoins de la communication, les locuteurs ont opéré un changement interne en utilisant DEF N A. La polydéfinitude DEF N DEF A a servi d'intermédiaire entre la monodéfinitude DEF A N et la monodéfinitude DEF N A. Elle a été favorisée par le contact de langues, plus précisément par l'existence d'un trait similaire dans les langues de contact principales que sont l'albanais et le grec. Le Tableau 112 résume ces hypothèses.

Tableau 112 : Chronologie de l'acquisition de l'article défini, de la polydéfinitude et de la postposition de l'adjectif dans les variétés balkaniques

Jadis :	**DEF A N**	**DEF A DEF N**	**DEF N DEF A**	**DEF N A**
Proto-romani	✗	✗	✗	✗
Vieux romani 1)	✓[1]	✗	✗	✗
2)	✓	✓[2]	✓[2]	✗
Aujourd'hui :				
romani de Grèce	✓	✓	✓	(✗)[3]
romani d'Albanie	✓	(✓)[4]	✓	✓[5]
romani de Yougoslavie, de Bulgarie	✓	(✓)[4]	✓	✗
romani de Roumanie, de Moldavie	(✗)[6]	(✓)[4]	✓	✓

Légende :

1. emprunt de l'article défini au grec byzantin
2. emprunt de la polydéfinitude au grec byzantin
3. structure marquée
4. structure devenue désuète
5. innovation interne
6. structure devenue marquée

4) Existe-t-elle dans d'autres variétés du romani ? Constate-t-on des différences ?

La polydéfinitude existe en romani d'autres pays que l'Albanie. On la trouve au minimum dans tous les autres pays des Balkans, notamment en Roumanie. Par suite des migrations des locuteurs balkaniques, on la trouve par exemple en Suède ou en Colombie, ou encore en France et au Royaume-Uni où nous l'avons entendue.

Il ne semble exister aucune différence entre les syntagmes polydéfinis des différentes variétés de romani. Nous avons constaté qu'en Roumanie, la polydéfinitude, la postposition de l'adjectif (en tous syntagmes et avec tous déterminants) et la surdétermination sont bien plus fréquentes qu'en romani d'Albanie. La réplication de la flexion du nom sur l'adjectif et de la préposition fusionnée avec l'article sont également pratiquées. La flexion sur l'article et sur l'adjectif est toutefois soumise à d'importantes variations phonétiques, voire n'est pas respectée.

5) La syntaxe du syntagme nominal est-elle stable en romani ?

La syntaxe nominale canonique veut que l'adjectif soit antéposé en romani. Toutes les variétés de romani ne connaissent pas la postposition de l'adjectif en syntagme monodéfini (DEF N A). Celles parlées dans des zones où les langues de contact antéposent l'adjectif ne l'ont probablement pas acquise et leur syntaxe nominale est donc stable.

Celles parlées dans des zones où les langues de contact postposent l'adjectif de façon massive (Italie, Roumanie) ont pu acquérir la postposition de l'adjectif, qu'elles aient ou non conservé la polydéfinitude. La polydéfinitude, si elle était encore en usage, a dû faciliter le processus. Leur syntaxe nominale est donc stable également. Prenons l'exemple de la Roumanie : nous avons relevé des occurrences de postposition de l'adjectif dans des textes roumains du XIXème siècle, qui présentent les mêmes caractéristiques syntaxiques que les variétés roumaines contemporaines. La postposition de l'adjectif y semble hégémonique, quel que soit le déterminant introducteur (article défini, possessif, démonstratif, article indéfini, article zéro, quantifieur). Cela montre que le changement de l'ordre des mots remonte à une époque reculée et qu'il est bien achevé.

6) La syntaxe du syntagme nominal est-elle stable en romani d'Albanie ?

La syntaxe nominale n'est pas stable en romani d'Albanie. À la lumière de nos résultats et de l'histoire du contact linguistique des différentes variétés, on constate qu'un changement de l'ordre des constituants nominaux s'est produit ou est en cours. L'ordre des mots initial est DEF A N. Il est utilisé en priorité par les locuteurs âgés et le groupe conservateur des Arlis. L'ordre des mots innovant est DEF N A. Il est utilisé en priorité par les locuteurs jeunes, les hommes et les groupes innovateurs des Mečkars et des Čergars I. L'ordre des mots DEF N DEF A représente un entre-deux, puisqu'il s'agit d'une rétention où l'adjectif est postposé. Il est utilisé en priorité par les locuteurs d'âge moyen et par les femmes.

À l'échelle de l'ensemble des variétés d'Albanie, on peut dire que DEF N A n'est plus une structure marquée. Si l'on considère l'usage de DEF N DEF A, où l'adjectif est postposé, on peut même dire

que la postposition est devenue majoritaire dans la langue, et la place non marquée de l'adjectif dans le contexte de nos tests. Il en va de même pour les syntagmes indéfinis.

Questions typologiques

1) La polydéfinitude existe-t-elle dans d'autres langues que le romani et le grec ?

La polydéfinitude existe, sous diverses formes, dans plusieurs langues des Balkans, ainsi qu'en langue familière en Europe centrale. Elle présente donc un caractère aréal. Elle présente également un caractère génétique, dans les langues germaniques septentrionales, les variétés de haut-allemand, les langues roumaines et langues sémitiques.

La postposition de l'adjectif présente quant à elle un caractère plus génétique. On la trouve dans les langues romanes et celtiques, ainsi qu'en maltais (sémitique) et en basque (isolat – cerné de langues romanes).

2) Quels sont les différents « types » de polydéfinitude dans les langues d'Europe ?

La définitude peut se marquer sur l'adjectif (lituanien) ou sur le nom (français). Elle peut consister en une marque flexionnelle adjectivale (serbo-croate) ou en un article défini (grec). L'article défini peut être libre (allemand *die*) ou lié (roumain *-ul*). La polydéfinitude peut se faire à l'aide de deux éléments identiques ou de deux éléments différents. Elle peut consister en la réplication de l'article défini libre, la réplication de l'article défini lié, ou en une combinaison de l'article libre et de l'article lié. Elle implique le plus souvent la postposition de l'adjectif.

La polydéfinitude peut revêtir un caractère « obligatoire » ou « optionnel » sur le plan grammatical. Dans le premier cas, elle ne présente aucun apport sémantique particulier. Dans le second cas, elle a un apport sémantique ou pragmatique. En romani, il n'y a pas nécessairement d'apport sémantique : si le syntagme est structurellement polydéfini, son sens reste monodéfini. Le Tableau 113 résume les différents types de marquage multiple de la définitude dans les langues d'Europe.

Tableau 113 : Typologie du marquage multiple de la définitude dans les langues d'Europe

	réplication		**combinaison mixte**
	de l'article défini libre	**de l'article défini lié**	
obligatoire	grec de Cappadoce	hébreu arabe	albanais suédois norvégien
optionnelle	grec standard romani slovène familier allemand familier bavarois yiddish	aroumain maltais	roumain, moldave aroumain, mégléno-roumain féroïen

3) Où se situe le romani dans cette typologie ?

Le romani correspond au type le plus fréquent de polydéfinitude : optionnelle, effectuée par réplication, d'un article défini libre, avec adjectif postposé. Il appartient à la catégorie des langues à *pattern* syntaxique, selon la typologie des déterminants multiples d'Alexiadou (2014 : 111-121).

Il contredit l'affirmation initiale de Lekakou et Szendrői (2007 : 152) selon laquelle « *we seem to find polydefinites of the Greek type only in Greek, even though appositives are attested elsewhere* », puisque le romani présente des syntagmes polydéfinis du type grec. Il confirme en revanche l'hypothèse typologique postérieure de Lekakou et Szendrői (2012 : section 5), selon laquelle seule une langue comportant à la fois un système de flexion nominale et la détermination obligatoire des noms propres avec un article défini explétif peut comporter des syntagmes polydéfinis.

4) Comment nommer au mieux la structure qui fait l'objet de la présente étude ?

Plusieurs noms existent pour ce phénomène : surdétermination (*overdetermination*), double détermination, double articulation, double définition, construction polydéfinie, syntagme nominal polydéfini, expansion/diffusion du déterminant (*determiner spreading*), répétition du déterminant, redoublement/réduplication du déterminant, déterminant/détermination multiple…

Nous pensons que les termes de *surdétermination* et de *déterminant/détermination multiple* ne sont pas adéquats, car ils vont au-delà de la structure qui nous intéresse. Ils incluent tous les cas de détermination multiple, par exemple les syntagmes :

- sans réplication de l'article défini ni modifieur adjectival : *le kodole amarë romes*, nos Roms/nos maris (structure DEF DEM POSS N), *sea-l boàle kadala*, toutes ces maladies (structure QUANT DEF N DEM), *me duj me phenja*, mes deux sœurs (POSS NUM POSS N)
- sans réplication de l'article défini avec modifieur adjectival : *amaro p<h>ral kou baro*, notre grand frère (POSS N DEM A), *muro pînro ko parno*, mon pied blanc (POSS N DEM A)
- avec réplication de l'article défini mais sans modifieur adjectival : *le kurùza le kodola*, ces céréales (DEF N DEF DEM), *o kodowa o ćhaw'*, ce garçon (DEF DEM DEF N)
- avec réplication de l'article défini et modifieur adjectival : *akova o roza o manyćyri*, ce vernis à ongles rose (DEM DEF A DEF N), *olenza i trinenza i muršenza*, avec ces trois hommes (DEM DEF NUM DEF N, analysable comme DEM DEF A DEF N), *ol thema ol strëine kodola*, ces pays étrangers (DEF N DEF A DEM), *e tradìcia amari e romaii*, notre tradition rom (DEF N POSS DEF A)

De plus, le terme *surdétermination* implique une idée d'excès. Les termes *déterminant multiple* et *détermination multiple* incluent jusqu'aux phénomènes de poly*in*définitude (Alexiadou 2014).

Le terme *double détermination* n'est pas adéquat, car il englobe les syntagmes à deux déterminants, qui sont très variés et fréquents en romani des Balkans (et dans les autres langues des Balkans). Nous avons relevé des cas où l'article défini accompagne le démonstratif (DEM DEF N,

DEF N DEM), le possessif (POSS DEF N, DEF N POSS), le quantifieur (QUANT DEF N, DEF N QUANT), le numéral (DEF NUM N, DEF NUM DEF N) et même l'article indéfini (DEF IND N)…

Les termes de *double articulation, double définition, construction polydéfinie, syntagme polydéfini* sont en-deçà de la structure qui nous intéresse. Ils signifient la présence de deux ou plusieurs articles définis, or cela englobe les appositions étroites nominales de structure DEF N DEF N (ou POSS N DEF N ou DEM N DEF N) :

- Albanie : *o gulipe i tatlija*, la pâtisserie tatli (*F53A, Korçë, 25 juillet 2014*)
- Serbie : *me phralesa e Marcelesa*, avec Marcel mon frère (*M63GSb, Zemun, 25 mars 2014*)
- Grèce : *ka o Jáni ka o Paríδi*, à Jani Paridi, *ka po dad ka o thagár*, à son père le roi (Igla 1996)
- romani de Roumanie *ka-o pòdo ka-o Nistro*, au pont du Dniestr (Cioabă 2006a)

De plus, les termes de *double définition, construction polydéfinie, SN polydéfini* ou même *marquage multiple de la définitude* sont-ils vraiment adéquats si l'article défini est explétif en romani, donc le syntagme n'est pas poly-défini sur le plan sémantique ?

Le terme *double articulation* implique quant à lui qu'il y ait deux articles définis, or ce n'est pas toujours le cas. Par exemple, l'article défini n'est pas redoublé dans les structures POSS N DEF A : *mo phral o xurdo*, mon petit frère. Les termes *redoublement/réduplication du déterminant* et *répétition du déterminant* ne conviennent pas pour la même raison : l'article défini n'est pas toujours redoublé.

Les termes d'*expansion du déterminant* (*determiner spreading*), *redoublement/réduplication du déterminant* et *répétition du déterminant* vont au-delà de notre sujet d'étude, car ils englobent les appositions nominales étroites. De plus, ils signifient toute réduplication de déterminant. Or en romani, le phénomène de réduplication se limite à l'article défini. Nous avons certes relevé quelques exceptions, mais elles sont trop peu nombreuses à l'échelle de l'ensemble de la langue : une occurrence de double démonstratif en romani d'Albanie contemporain (notre corpus semi-spontané), une en romani roumain ancien (recueil de Constantinescu 2016), une en romani de Serbie (corpus RMS) et deux occurrences en romani de Grèce (Matras 2004 : 75-76).

Par conséquent, aucun terme ne nous semble vraiment satisfaisant. De plus, les termes existants nous semblent faire l'impasse sur la présence de l'adjectif et ses caractéristiques. Le présent travail porte sur les **syntagmes nominaux** comportant un **adjectif épithète** accompagné d'un **article défini** et en général **postposé** au nom. Ils s'opposent à ceux qui ne comportent pas d'adjectif, à ceux où l'adjectif n'est pas épithète du nom-tête, n'est pas accompagné d'un article défini (syntagmes que nous avons baptisés « monodéfinis ») ou est antéposé au nom. Nous préférons conserver le terme de *syntagme nominal polydéfini*, malgré toutes les réserves que nous avons émises, notamment quant au concept de polydéfinitude.

Figure 87 : Fin de la réponse à la problématique

Conclusion

Cette thèse a pour objet l'analyse des syntagmes nominaux polydéfinis en romani d'Albanie. « Polydéfini » signifie qu'un article défini supplémentaire précède l'adjectif dans un syntagme nominal déjà introduit par un article défini. On parle également de polydéfinitude, d'expansion du déterminant ou d'articulation multiple. Cet article « adjectival » n'est pas requis par la grammaire mais est optionnel, comme c'est le cas dans la structure similaire du grec moderne.

Cette thèse a permis de montrer qu'il fait l'objet d'une variation complexe, à la fois dialectale, diachronique, diagénique et diaphasique. On trouve cette structure dans la plupart des variétés balkaniques du romani. En Albanie, où nous avons réalisé nos missions de terrain, les personnes qui l'emploient le plus sont les Mečkars et les Čergars I, les locuteurs d'âge moyen et les femmes, pour des raisons essentiellement pragmatiques.

Méthodologie

Le travail par jugement d'acceptabilité et par élicitation ne s'est pas révélé fructueux en ce qui concerne notre objet d'étude et la langue romani. Nous sommes partie du postulat que toute méthode comporte des biais et des inconvénients qui peuvent être compensés par une autre méthode. L'étude de corpus annoté de données spontanées et une expérience psycho-linguistique produisant des données semi-spontanées sont complémentaires. C'est pourquoi nous avons croisé des types de corpus et d'analyse différents, en combinant les méthodes quantitatives et qualitatives. Afin de vérifier les hypothèses de départ, nous avons effectué des recoupements d'information afin de vérifier leur véracité. Cela a permis de mettre à jour des « contraintes préférentielles » qui agissent dans les zones de liberté offertes par la syntaxe de la langue, selon qu'en donne Thuilier (2015 : 33).

Description de la structure

Le syntagme polydéfini est constitué d'un nom précédé d'un article défini et d'un adjectif lui aussi précédé d'un article défini. Il ne semble pas exister de restriction quant au type de nom entrant dans un syntagme polydéfini : nom propre ou commun, dénombrable ou indénombrable, abstrait, événementiel, animé ou inanimé. Il peut être hérité ou emprunté, de poids léger ou lourd.

Nous n'avons pas trouvé de contrainte de morphologie, de poids ni d'origine étymologique quant au type d'adjectif. Il semble exister une contrainte syntaxique, puisque nos corpus ne comprennent que des adjectifs prédicatifs. Sur le plan sémantique, les données recueillies présentent une majorité d'adjectifs orientés vers l'objet, restrictifs et intersectifs, qui permettent une interprétation conjointe du syntagme nominal. Lorsque l'adjectif n'est pas restrictif, il sert à « mettre une étiquette » (*epithet*) – phénomène observé dans les syntagmes polydéfinis grecs. Les adjectifs relationnels ou lexicalisés avec le nom sont rares mais possibles, s'ils impliquent un contraste implicite ou explicite.

Il existe des contraintes quant au type de déterminant. Il doit être nécessairement défini (simple ou complexe). Pas plus qu'en grec, il n'existe de structure poly*in*définie. Si le syntagme polydéfini « canonique » contient deux articles définis, nous constatons que le premier déterminant peut en réalité être un autre déterminant défini : possessif (POSS N DEF A) ou démonstratif (DEM N DEF A). Le premier déterminant peut également être « double » : démonstratif et article défini (DEM DEF N DEF A), article défini et quantifieur (DEF QUANT N/ DEF A). En romani de Roumanie, nous avons relevé des cas de triple détermination, avec réduplication de l'article défini associé à un démonstratif (DEF N DEM DEF A, DEF N DEF A DEM) ou à un possessif (DEF N POSS DEF A), ou encore article défini, possessif et numéral (POSS N DEF NUM A, POSS NUM N DEF A). Il existe également des cas de quadruple détermination avec réduplication de l'article défini associé au possessif et au numéral (DEF NUM DEF A POSS N) ou réduplication du possessif associé à un article défini et à un numéral (POSS NUM POSS N DEF A). Il n'y a donc pas à proprement parler de *poly-définitude*, en ce sens qu'il n'y a pas toujours plusieurs instances de l'article défini : mais il y a bien plusieurs instances d'un déterminant défini. Notons quelques hapax : redoublement du démonstratif, DEM N DEM A (un cas dans notre corpus albanais et un cas dans un texte roumain ancien), redoublement du possessif, POSS N POSS A ou POSS A POSS N (un cas dans un corpus serbe et deux cas dans une publication sur le romani grec).

Comme en grec, la polydéfinitude donne lieu à la répétition de la préposition lorsqu'elle a fusionné avec l'article défini. Si un syntagme polydéfini est introduit par une préposition monosyllabique cliticisée à DEF_1, on la trouve cliticisée à DEF_2, donnant lieu à une structure PREP.DEF_1 N PREP.DEF_2 A. Elle donne également lieu à la réplication des marques flexionnelles : la marque de flexion sur le nom est répétée sur l'adjectif.

Syntaxe

L'adjectif est presque toujours postposé au nom (DEF N DEF A) en syntagme polydéfini : on trouve très peu de syntagmes DEF A DEF N. Le romani se distingue donc là du grec. La consultation de données en romani albanais ancien, roumain ancien et en romani d'autres régions confirme ce constat. L'ordre des mots est donc moins souple en syntagme polydéfini que monodéfini. En effet, nous avons constaté qu'en syntagme monodéfini, le romani albanais permet l'antéposition de l'adjectif (DEF A N) aussi bien que sa postposition (DEF N A) en syntagme monodéfini. Il se distingue là du grec qui ne permet que l'antéposition de l'adjectif (DEF A N, *DEF N A).

Il est rare, mais possible, que le syntagme comporte plusieurs adjectifs épithètes : ils encadrent alors le nom (DEF A N DEF A) ou bien sont tous deux postposés. Dans ce deuxième cas, soit le premier est articulé (DEF N DEF A A), soit le deuxième (DEF N A DEF A), soit tous les deux (DEF N DEF A DEF A). La structure est donc récursive, mais de manière non systématique.

Le syntagme polydéfini est une apposition étroite, comprenant une partie nominale et une partie adjectivale qui constituent une seule unité prosodique, portant un seul accent. Les syntagmes polydéfinis avec rupture ressemblent à l'apposition lâche décrite pour le grec ; ils comprennent deux sous-unités prosodiques et deux accents nucléaires. La rupture entre sous-partie nominale et adjectivale consiste en une pause, parfois en une insertion lexicale (nom propre, interjection, proposition relative, voire le reste de l'énoncé). Pour autant, il ne s'agit pas de deux syntagmes juxtaposés DEF N DEF N où l'adjectif est substantivé. En effet, lorsque l'adjectif substantivé est

introduit par un article défini, il n'est alors pas restrictif. Lorsque l'adjectif substantivé est restrictif, il n'est alors pas introduit par un article défini, mais par un démonstratif.

L'ordre des constituants nominaux

L'emplacement post-nominal du syntagme permet un emploi spécifique et classificateur de l'adjectif. Il est le lieu préférentiel d'emploi des adjectifs empruntés, morphologiquement adaptés ou non, empruntés à une langue antéposant ou postposant l'adjectif et quelle que soit l'époque de l'emprunt. Il a donc une fonction dans la compartimentation de la syntaxe nominale du romani albanais. De plus, le focus interne au syntagme nominal réside dans un emplacement syntaxique donné, et c'est l'emplacement post-nominal. Il permet un focus étroit sur l'adjectif postposé, probablement aussi sur le syntagme génitival postposé, ce que ne permet pas l'antéposition. Il a pour fonction le contraste, ce qui explique que la postposition de l'adjectif soit non-marquée dans le contexte de la tâche, où l'adjectif est restrictif. Cela s'explique aussi par le fait que grâce au focus, les sèmes spécifiques de l'adjectif sont mis en avant. De même, lorsque plusieurs adjectifs encadrent le nom, c'est l'adjectif postposé qui porte le contraste. Comme l'adjectif restrictif est nécessaire à l'identification, il n'est nullement une option.

L'emplacement postnominal constitue la « post-zone » au sens de Nølke (2001), c'est-à-dire qu'elle est le lieu du focus adjectival, contrairement à la « pré-zone ». Nos analyses du romani albanais montrent qu'il n'est pas seulement la zone des « options » telle que définie par Matras (2002 : 166). On y trouve des modifieurs épithètes : adjectifs non-articulés, adjectifs articulés, noms en apposition étroite, syntagmes génitivaux. On y trouve également des déterminants (article défini, démonstratif, possessif, numéral) qui clôturent le syntagme ou qui accompagnent l'adjectif postposé ou le nom en apposition étroite. Cela n'empêche pas d'y trouver également des « options » ayant la fonction d'*afterthought*, ainsi que le montre Matras (2002 : 166) pour d'autres variétés de romani. Le romani albanais présente selon nous la même topologie syntaxique et sémantique que les langues romanes selon Cinque (2010, 2014) : la position antéposée est celle de la modification directe, la position postposée est celle de la modification directe ou indirecte. Pour nos corpus, nous avons analysé la partie adjectivale des syntagmes polydéfinis (avec et sans rupture interne) comme des cas de modification indirecte.

Le recensement des syntagmes polydéfinis en romani d'Albanie, d'autres régions et d'autres époques a montré qu'il existe un lien avec entre polydéfinitude et ordre des constituants nominaux. Contrairement au grec, la polydéfinitude du romani se manifeste principalement lorsque l'adjectif est postposé. Toutes les variétés qui comportent des syntagmes polydéfinis peuvent postposer l'adjectif en syntagme monodéfini, dans un certain but communicatif et de façon marquée ou non. Or on sait que les relations syntaxiques font volontiers l'objet d'un transfert linguistique : la postposition adjectivale, probablement issue d'une innovation interne dans certaines variétés, a été favorisée par le contact de langues, notamment pour les variétés en contact avec des langues postposant l'adjectif. Aujourd'hui, le placement de l'adjectif est régi par des contraintes préférentielles et pragmatiques.

Sémantique

En syntagme monodéfini comme en syntagme polydéfini, l'interprétation est soit restrictive soit non-restrictive. Il existe toutefois une nuance de sens : les syntagmes polydéfinis sont plus fréquemment restrictifs, partant plus spécifiques. *Poly*définitude n'implique pas *sur*définitude : le syntagme n'est pas « plus » défini sur le plan sémantique, car le deuxième article défini n'a pas pour fonction de rendre plus défini le syntagme. Le syntagme polydéfini a pour caractéristique d'effectuer une restriction pragmatique parmi les référents possibles du nom. La fonction majeure de cette structure est l'identification. Elle est employée lorsque le nom dénote plus de deux référents possibles « concurrents » et que l'identification peut être considérée comme délicate. Dans les cas (beaucoup moins nombreux) où le syntagme a un effet d'emphase, cette fonction d'identification ne semble pas être présente. Puisque la redondance de l'article défini contribue à l'interprétation pragmatique du syntagme, ce phénomène n'est pas un cas de redoublement syntaxique (*syntactic doubling*).

La position de l'adjectif joue un rôle important dans l'interprétation. Lorsque l'adjectif est antéposé, l'interprétation est plutôt non-restrictive ; lorsque l'adjectif est postposé, elle est plutôt restrictive. En syntagme monodéfini, lorsque l'adjectif est antéposé, l'interprétation est restrictive ou non-restrictive ; lorsque l'adjectif est postposé, elle est plutôt restrictive.

Notons que nous ne sommes pas parvenue à opérer de distinction sémantique nette entre syntagmes polydéfinis sans rupture interne et avec. Certes, les syntagmes avec rupture sont moins systématiquement restrictifs, mais ils peuvent l'être : cela constitue une différence avec le grec, où un syntagme avec pause ou insertion implique que l'adjectif apporte une information non essentielle et non-restrictive.

Structure et statut de l'information

Les syntagmes polydéfinis font partie soit du topique, soit du focus de la phrase. Ils transmettent une information nouvelle, donnée ou accessible. L'adjectif, lui, est toujours nouveau lexicalement : il apporte une propriété nouvelle qui permet l'identification du référent.

Le syntagme polydéfini comporte un focus étroit sur sa partie adjectivale, de type informatif ou contrastif, explicite ou implicite. L'effet de sélection va de pair avec un contraste *in presentia* (référents du même ensemble mentionnés en discours ou présents dans la situation d'énonciation) ; l'effet d'emphase va de pair avec un contraste *in absentia*. Le premier, *in presentia*, qui correspond à l'effet de sélection explicite, se trouve plus souvent dans les syntagmes avec que sans rupture. L'article défini « adjectival » permet de focaliser l'adjectif. Il joue le rôle d'un opérateur d'identification ou de focalisation, qui renforce le statut référentiel de cette partie et de l'ensemble du syntagme. La présence d'un article défini devant l'adjectif augmente le statut informatif de celui-ci, de même que la postposition : l'adjectif antéposé au nom ne porte pas de focus propre, contrairement à l'adjectif postposé. Nous sommes arrivée à la conclusion que la focalisation n'est pas un phénomène binaire mais scalaire, et que les syntagmes nominaux, selon leur structure (monodéfinie, polydéfinie, à adjectif antéposé ou postposé) se répartissent le long de cette échelle. Nous pensons également que l'article défini « adjectival » n'est pas tout à fait un déterminant, dans la mesure où il ne satisfait pas les contraintes d'unicité, d'existence ni de familiarité. Il serait donc explétif.

Il apparaît que la structure informationnelle du syntagme nominal est déterminée par sa structure ou la « géographie » interne. La post-zone est le lieu du focus, la pré-zone est le lieu, non-focal, de la présupposition. Du moins est-ce le cas dans les variétés mečkar et čergar I, l'arli présentant actuellement une distinction moins nette entre les deux.

Fonctions

Nous sommes convaincue que tout, dans la langue, a une fonction. La structure sert un focus contrastif consistant à sélectionner un référent au sein d'un ensemble de référents similaires (fonction d'identification), ainsi qu'un focus assertif permettant un effet d'emphase sur l'adjectif. Elle a également une fonction de précision du référent déjà mentionné au préalable, d'auto-correction. Elle a aussi une fonction de récapitulation du discours qui précède ainsi qu'une fonction de *topic-shift*. Elle permet la création de noms propres familiers et de surnoms. Enfin, en romani de Roumanie, elle est perçue comme une structure « héritée » et sert à « romaniser » des expressions ou concepts empruntés à la langue de contact.

Comme en grec, le syntagme polydéfini est optionnel et fait l'objet d'un choix du locuteur. Celui-ci est contraint par la modification adjectivale, le contexte défini, la présence d'un contraste implicite ou explicite, la postposition de l'adjectif, l'emploi restrictif ou d'« étiquette » de l'adjectif. L'usage de la polydéfinitude marque la prise en compte de la relation interlocutive. Il indique que le locuteur se préoccupe de la manière dont son interlocuteur pensera le référent. Au contraire, l'usage de la monodéfinitude révèle que le locuteur n'a pas conscience de la présence d'éventuels référents « concurrents » dans la dénotation du nom.

Réalité socio-linguistique

Tous les locuteurs albanais connaissent la structure polydéfinie, quels que soit leur âge, genre, niveau d'éducation et variété, qu'ils soient ou non militants. L'analyse a montré que les variables MILITANTISME et EDUCATION sont moins pertinentes que les autres pour expliquer l'usage de la polydéfinitude, contrairement à celle du GENRE, puisque l'on constate que les femmes emploient plus la structure que les hommes. Les deux variables les plus importantes sont AGE et VARIETE : la structure est pratiquée le plus volontiers par les adultes ayant entre 15 et 59 ans, et par les locuteurs du mečkar et du čergar I.

La personne qui a le plus de probabilité de produire fréquemment DEF N DEF A est une femme, non-militante, âgée entre 15 et 59 ans et parlant mečkar ou čergar I. Le profil le plus fréquent pour DEF N A est un homme de moins de 15 ans, en cours de scolarisation (ou moyennement scolarisé s'il est plus âgé), non-militant, parlant čergar I (mono- ou bi-variétal) ou mečkar. Le profil le plus fréquent pour DEF A N est un homme ou une femme de plus de 60 ans, militant, qui parle arli (monovariétal, bivariétal ou néo-locuteur).

À la lumière de nos résultats et de l'histoire du contact linguistique des différentes variétés, nous avons constaté qu'un changement de l'ordre des mots dans le syntagme nominal s'est produit ou est en cours. L'ordre des mots initial est DEF A N. Il est utilisé en priorité par les locuteurs âgés et le groupe

syntaxiquement conservateur des Arlis. L'ordre des mots innovant est DEF N A. Il est utilisé en priorité par les locuteurs jeunes, les hommes et les groupes innovateurs des Mečkars et des Čergars I. L'ordre des mots DEF N DEF A représente un entre-deux, puisqu'il s'agit d'une rétention où l'adjectif est postposé. Il est utilisé en priorité par les locuteurs d'âge moyen et par les femmes. En corrélant les variables les plus importantes, AGE et VARIETE, on distingue les groupes suivants :

- Les Mečkars et Čergars I connaissent une phase de *stabilité* linguistique, car ils ont achevé le changement de l'ordre initial DEF A N vers DEF N DEF A puis l'ordre innovateur DEF N A.
- Les Arlis âgés, conservateurs sur le plan syntaxique, connaissent également la *stabilité* car ils n'ont pas opéré de changement.
- Les Arlis d'âge moyen expérimentent le modèle du *changement au cours de la vie* car ils ont adopté, au cours de leur vie, le changement en question, mais sans l'avoir systématisé.
- Les jeunes Arlis emploient massivement la variante innovante, ce qui peut correspondre au modèle de la *gradation par l'âge* ou à celui du *changement générationnel*.

À l'échelle de l'ensemble des variétés d'Albanie, on peut dire que DEF N A n'est plus une structure marquée. Si l'on considère l'usage de DEF N DEF A outre celui de DEF N A, on peut même affirmer que la postposition est devenue majoritaire dans la langue, et la place non marquée de l'adjectif dans le contexte de nos tests. Il en va de même pour les syntagmes indéfinis.

Histoire

Des syntagmes polydéfinis sont attestés dans des textes en romani albanais des années 1930 et dans des textes en romani roumain du milieu du XIXème siècle. Nous pensons que cette structure remonte au vieux romani, donc à l'arrivée des Roms en Europe et à l'époque byzantine (XIème siècle ?), auquel cas toutes les variétés actuelles ont un jour connu la polydéfinitude dans leur histoire. Les Roms se sont ensuite dispersés et certaines communautés ont perdu l'usage de la polydéfinitude voire de l'article défini. Les variétés des zones où les langues de contact ne présentent pas d'article défini ont commencé par perdre la polydéfinitude, avant de perdre progressivement l'article défini lui-même. Celles des zones où les langues de contact présentent un article défini mais pas de polydéfinitude ont perdu l'usage de la polydéfinitude mais ont conservé l'article défini. Aujourd'hui, la polydéfinitude est donc utilisée de manière plus fréquente par les locuteurs bilingues d'une langue possédant un article défini et des syntagmes présentant un article défini adjectival. Si la structure provient du grec, d'autres langues présentant un article adjectival, comme l'albanais, ont pu favoriser son maintien ou son développement en romani. L'acquisition des syntagmes polydéfinis est un transfert linguistique par contact de langues, ainsi qu'une convergence de type *pattern replication* (Matras et Sakel 2007, Matras 2009). Les traits pivots sont la réduplication totale de l'article défini et la postposition de l'adjectif ; la langue R (réplique) est le romani, la langue M (modèle) le grec. Cette réplication grammaticale a provoqué un changement à long terme pour le romani, en enrichissant son inventaire structural des syntagmes DEF N DEF A et DEF A DEF N.

Le romani des Balkans pratique la postposition de l'adjectif depuis une époque très ancienne, postérieure à celle du vieux romani mais antérieure au XIXème siècle. En romani d'Albanie, l'acquisition de la postposition de l'adjectif (en syntagme monodéfini, polydéfini ou indéfini) est postérieure à l'acquisition de la polydéfinitude (à adjectif antéposé ou postposé). Le vieux romani,

possédant déjà DEF A N (hérité), a emprunté à la fois DEF A DEF N et DEF N DEF A au grec. Le romani balkanique a ensuite progressivement perdu DEF A DEF N dans les zones où la langue de contact ne le pratique pas (Roumanie, Albanie). Nous pouvons supposer qu'en romani albanais, la structure canonique DEF A N a cessé d'être adéquate dans certains contextes. Pour mieux répondre aux besoins de la communication, les locuteurs ont opéré un changement interne en utilisant DEF N A. La polydéfinitude DEF N DEF A a servi d'intermédiaire entre la monodéfinitude DEF A N et DEF N A. Il s'agit d'une innovation interne, probablement favorisée par le contact de langues, plus précisément par l'existence d'un trait similaire dans les langues de contact principales que sont l'albanais et le grec. C'est un cas de réplication grammaticale affectant les relations syntaxiques, effectuée avec la stratégie du *shift from one construction type to another* (Heine 2008). Il s'agit également d'un cas de convergence linguistique, de type *pattern replication* où le trait pivot est la postposition de l'adjectif (de classes 1 et 2) côté albanais, et la postposition de l'adjectif en syntagme polydéfini côté romani. La langue R est le romani, la langue M est l'albanais, le roumain, ou toute autre langue de contact locale qui postpose l'adjectif. Cette *pattern replication* de l'albanais a été rendue possible par celle faite du grec, c'est-à-dire que DEF N DEF A a permis l'acquisition de DEF N A, en servant d'intermédiaire entre DEF A N et DEF N A. En effet, la structure DEF N DEF A est très ancienne, tout comme DEF A N, et postpose l'adjectif, tout comme DEF N A. La généralisation de la postposition de l'adjectif en syntagme monodéfini par certaines variétés relève du *pragmatic unmarking*.

Géographie

Les syntagmes polydéfinis existent en romani d'Albanie, mais aussi dans les autres pays des Balkans, notamment en Roumanie, ainsi que dans les pays de migration tels que la Suède, la Colombie, ou encore en France et au Royaume-Uni[1]. Il ne semble exister aucune différence d'emploi d'une variété à l'autre. Tout au plus trouve-t-on des différences de fréquence et de morphologie. En Roumanie, la polydéfinitude, la surdétermination et la postposition de l'adjectif (en tous syntagmes et avec tous déterminants) sont bien plus fréquentes qu'en Albanie. La flexion de l'adjectif et de son article est soumise à fluctuation et à d'importantes variations phonétiques.

La postposition de l'adjectif monodéfini existe en Albanie, mais aussi dans d'autres variétés des Balkans et d'Europe. Celles parlées dans des zones où les langues de contact postposent l'adjectif de façon massive (Italie, Roumanie) ont pu acquérir la postposition de l'adjectif, qu'elles aient ou non conservé la polydéfinitude (qui, si elle était encore en usage, a facilité le processus). En revanche, celles parlées hors des Balkans et dans des zones où les langues de contact antéposent l'adjectif ne l'ont probablement pas acquise.

Typologie

La polydéfinitude se rencontre sous diverses formes dans plusieurs langues des Balkans et d'Europe centrale : elle présente donc un caractère aréal. Elle présente également un caractère génétique, en ce qui concerne les langues germaniques septentrionales, les variétés de haut-allemand, les langues

[1] où nous en avons entendu des occurrences lors de conversations de la vie courante avec des Roms roumains.

roumaines et langues sémitiques. La postposition de l'adjectif présente quant à elle un caractère essentiellement génétique. On la trouve dans les langues romanes et celtiques, ainsi qu'en maltais (sémitique) et en basque (isolat – cerné de langues romanes).

Qu'entend-on par « diverses formes » ? Il peut s'agir de deux articles identiques ou différents. Elle peut consister en la réplication de l'article défini libre, la réplication de l'article défini lié, ou en une combinaison d'article libre et lié. Nous avons remarqué qu'elle implique le plus souvent la postposition de l'adjectif. La polydéfinitude peut revêtir un caractère « obligatoire » ou « optionnel » sur le plan grammatical. Dans le premier cas, elle ne présente aucun apport sémantique particulier. Dans le second cas, elle a un apport sémantique ou pragmatique. Le romani correspond au type le plus fréquent de polydéfinitude : optionnelle, effectuée par réplication, d'un article défini libre, avec adjectif postposé. Il appartient à la catégorie des langues à *pattern* syntaxique, selon la typologie des déterminants multiples d'Alexiadou (2014 : 111). Contrairement à l'affirmation initiale de Lekakou et Szendrői (2007 : 152) selon laquelle « *we seem to find polydefinites of the Greek type only in Greek, even though appositives are attested elsewhere* », la polydéfinitude du type grec n'existe pas ailleurs qu'en grec. Le romani confirme en revanche l'hypothèse typologique postérieure de Lekakou et Szendrői (2012 : section 5), selon laquelle seule une langue comportant à la fois un système de flexion nominale et la détermination obligatoire des noms propres avec un article défini explétif peut comporter des syntagmes polydéfinis.

Perspectives

Il serait pertinent de compléter ce travail par une étude phonologique. Il s'agirait d'étudier la prosodie dans les énoncés comportant des syntagmes polydéfinis et dans les syntagmes eux-mêmes. Cela nous aiderait peut-être à percer le mystère des syntagmes polydéfinis avec rupture interne, en les comparant aux syntagmes sans rupture, puisque nous n'avons pas décelé de différence sémantique ou fonctionnelle. Cela permettrait de déterminer s'il existe des syntagmes polydéfinis lâches en romani, distincts des syntagmes polydéfinis étroits. Une étude de l'emplacement du syntagme génitival en romani d'Albanie confirmera nos conclusions concernant le rôle informationnel et sémantique de la post-zone dans le syntagme nominal.

Afin de confirmer nos conclusions diachroniques et géographiques, nous souhaiterions également étudier le romani de Suède et de Norvège, afin d'observer s'il comporte des syntagmes polydéfinis, et quelle est la place de l'adjectif dans ces variétés. Nos prédictions typologiques prévoient que ces variétés aient pu conserver la polydéfinitude, mais utiliser peu la postpostion de l'adjectif – comme c'est le cas en romani de Grèce. Enfin, nous souhaitons observer la postposition de l'adjectif dans les variétés du romani en contact avec des langues antéposant l'adjectif (serbe, bulgare…) ou le postposant (italien, français, catalan). Il faudrait comprendre quels facteurs déterminent la place de l'adjectif, dans quels dialectes. On prendrait en considération la nature du déterminant qui introduit les syntagmes concernés. Ces études permettraient de poursuivre le travail engagé et de l'élargir.

Bibliographie

Abeillé, Anne et Danielle Godard, 1999. « La position de l'adjectif épithète en français : le poids des mots ». *Recherches Linguistiques de Vincennes* 28 : 9-32.

Ackerley, Frederick G. et Eric O. Winstedt, 1925. « Specimens of Finnish Romani ». *Journal of the Gypsy Lore Society* III 4(2) : 73-85.

Acuña, Esteban et Evangelia Adamou, 2013. « How Did Linguistic Innovations Spread in the Romani Varieties of the Americas? A Social Interconnection Approach ». *Gypsy Lore Society Annual Meeting*. University of Strathclyde, Glasgow.

Ad Hoc Committee of Experts on Roma Issues (CAHROM), 2012. *Estimates and Official Numbers of Roma in Europe. Council of Europe.* http://rm.coe.int/CoERMPublicCommonSearchServices/DisplayDCTMContent?documentId=09000001680088ea9

Adamou, Evangelia, 2006a. « The Louse and the Rom ». *La Collection Pangloss*. http://lacito.vjf.cnrs.fr/pangloss/corpus/show_text.php?id=cocoon-b48c740d-59ad-3d91-9956-fdd3ea003ba0&idref=cocoon-1330a853-7d11-346b-97c2-cd52ec8b0231

———, 2006b. « The Coward and the Giants ». *La Collection Pangloss*. http://lacito.vjf.cnrs.fr/pangloss/corpus/show_text.php?id=cocoon-af3f71a6-b342-38e0-9707-ad82d5175fac&idref=cocoon-f4d1f3b5-7464-35d7-a716-d423146cf073

———, 2010. « Bilingual Speech and Language Ecology in Greek Thrace: Romani and Pomak in Contact with Turkish ». *Language in Society* 39(2) : 147-171.

———, 2011. « Temporal Uses of Definite Articles and Demonstratives in Pomak (Slavic, Greece) ». *Lingua* 121(5) : 871-889.

———, 2012. *Description, typologie et contact de langues : langues minoritaires des Balkans et du Mexique*. Habilitation à diriger des recherches, Paris : Institut National des Langues et Civilisations Orientales.

———, 2013. « Replicating Spanish *Estar* in Mexican Romani ». *Linguistics* 51(6) : 1075-1105.

———, 2015. « A Corpus-Driven Analysis of Romani in Contact with Turkish and Greek ». *Language Variation - European Perspectives V Selected Papers from the Seventh International Conference on Language Variation in Europe (ICLaVE 7)*, édité par Eivind Torgersen, Stian Hårstad, Brit Mæhlum et Unn Røyneland : 1-16.

———, 2016. *A Corpus-Driven Approach to Language Contact: Endangered Languages in a Comparative Perspective*. Berlin : Mouton de Gruyter.

Adamou, Evangelia et Amalia Arvaniti, 2011. « Focus Expression in Romani ». *84th Annual Meeting of the Linguistic Society of America (LSA) 2010*, édité par Mary Byram Washburn, Katherine McKinney-Bock, Erika Varis, Ann Sawyer et Barbara Tomaszewicz. 240-248. Los Angeles : Cascadilla Proceedings Project.

———, 2014. « Greek Thrace Xoraxane Romane ». *Journal of the International Phonetic Association* 44(2) : 223-231.

Adamou, Evangelia, Stefano De Pascale, Yekatarina Garcia-Markina et Cristian Pădure. 2019. « Do bilinguals generalize estar more than monolinguals and what is the role of conceptual transfer? » *International Journal of Bilingualism* 23 (6): 1549–1580.

Adamou, Evangelia et Kimmo Granqvist, 2015. « Unevenly Mixed Romani Languages ». *International Journal of Bilingualism* : 525-547.

Advisory Committee on the Framework Convention for the Protection of National Minorities, 2012. *Third Opinion on Albania Adopted on 23. November 2011*. Conseil de l'Europe. https://rm.coe.int/CoERMPublicCommonSearchServices/DisplayDCTMContent?documentId=090000168008c633

Alexiadou, Artemis, 2001. *Functional Structure in Nominals: Nominalization and Ergativity*. Amsterdam & Philadelphia : John Benjamins Publishing Company.

———, 2005. « Possessors and (in)Definiteness ». *Lingua* 115 : 787-819.

———, 2014. *Multiple Determiners and the Structure of DPs*. Amsterdam & Philadelphia : John Benjamins Publishing Company.

Alexiadou, Artemis et Elena Anagnostopoulou, 2000. « Greek Syntax: A Principles and Parameters Perspective ». *Journal of Greek Linguistics* 1 : 169-221.

Alexiadou, Artemis, Geoffrey Horrocks et Melita Stavrou, 1999. « Introduction ». *Studies in Greek Syntax. 1-21. Studies in Natural Language and Linguistic Theory*, édité par Artemis Alexiadou, Geoffrey Horrocks et Melita Stavrou. Dordrecht, Boston & London : Kluwer Academic Publishers.

Alexiadou, Artemis et Chris Wilder, 1998. « Adjectival Modification and Multiple Determiners ». *Possessors, Predicates and Movement in the DP*, édité par Artemis Alexiadou et Chris Wilder. 303-332. Amsterdam & Philadelphia : John Benjamins Publishing Company.

Aljović, Nadira, 2010. « Syntactic Positions of Attributive Adjectives ». *Adjectives: Formal Analyses in Syntax and Semantics*, édité par Patricia Cabredo Hofherr et Ora Matushansky. 29-52. Amsterdam & Philadelphia : John Benjamins Publishing Company.

Anagnostopoulou, Elena, 2006. « Clitic Doubling ». *The Blackwell Companion to Syntax*, édité par Martin Everaert et Henk van Riemsdijk. Oxford : Blackwell Publising.

———, 2013. « Greek Syntax 2012: State-of-the-Art and Perspectives ». *Proceedings of the 33rd Annual Meeting of the Department of Linguistics*. 11-58.

Andoni, Doris et Aida Orgocka, 2013. *Housing Policies and Practice for Roma in Albania*. Tiranë : Programme des Nations Unies pour le développement.

Androutsopoulos, Jannis, éd. 2014. *Mediatization and Sociolinguistic Change*. Berlin & Boston : Mouton de Gruyter.

Androutsopoulou, Antonia, 1995. « The Licensing of Adjectival Modification ». *Proceedings of West Coast Conference on Formal Linguistics 14*. 17-31. Stanford : CSL1.

———, 2001. « Adjectival Determiners in Albanian and Greek ». *Comparative Syntax of Balkan Languages*, édité par Maria-Luisa Rivero et Angeliki Ralli. 161-199. Oxford : Oxford University Press.

Ariel, Mira, 1988. « Referring and Accessibility ». *Journal of Linguistics* 24 : 65-87.

———, 1991. « The Function of Accessibility in a Theory of Grammar ». *Journal of Pragmatics* 16 : 443-463.

Assenova, Petya, 2002. *Balkansko ezikoznanie [Linguistique balkanique]*. Sofia : Faber.

Asséo, Henriette, 2003. « Les Tsiganes dans les Balkans ». *Matériaux pour l'histoire de notre temps* 71(1) : 30-38.

Avgustinova, Tania, 1998. « Determinedness and Replication Potential of Nominal Material in Bulgarian ». *Balkanistika* 11 : 1-17.

Babby, Leonard H., 2010. « The Syntactic Differences between Long and Short Forms of Russian Adjectives ». *Adjectives: Formal Analyses in Syntax and Semantics*, édité par Patricia Cabredo Hofherr et Ora Matushansky. 53-84. Amsterdam & Philadelphia : John Benjamins Publishing Company.

Baker, Mark C., 2003. *Lexical Categories: Verbs, Nouns, and Adjectives*. Cambridge & New York : Cambridge University Press.

Bakker, Peter, 1999. « The Northern Branch of Romani: Mixed and Non-Mixed Varieties ». *Die Sprache der Roma: Perspektiven der Romani-Forschung in Österreich im interdisziplinären und internationalen Kontext*, édité par Dieter W. Halwachs et Florian Menz. 172-209. Klagenfurt : Drava.

———, 2001. « Romani in Europe ». *The Other Languages of Europe: Demographic, Sociolinguistic and Educational Perspectives*, édité par Guus Extra et Durk Gorter. 293-313. Clevedon : Multilingual Matters.

———, 2012. « Romani Genetic Linguistics and Genetics: Results, Prospects and Problems ». *Romani Studies* 22(2) : 91-111.

Bakker, Peter, Milena Hübschmannová, Valdemar Kalinin et al., 2000. *What is the Romani Language?* Paris & Hatfield : Centre de recherches tsiganes & University of Hertfordshire Press.

Ball, Martin J. et Nicole Müller, 2010. *The Celtic Languages*. 2ème édition. London & New York : Routledge.

Banfi, Emanuele, 1985. *Linguistica Balcanica*. Bologna : Zanichelli.

Barbiers, Sjef, 2013. « Where is Syntactic Variation? ». *Language Variation - European Perspectives IV: Selected Papers from the Sixth International Conference on Language Variation in Europe (ICLaVE 6)*, édité par Peter Auer, Javier Caro Reina et Göz Kaufmann. 1-26. Amsterdam & Philadelphia : John Benjamins Publishing Company.

Barbiers, Sjef, Olaf Koeneman et Marika Lekakou, 2009. « Syntactic Doubling and the Structure of Wh-Chains ». *Journal of Linguistics* 1(46) : 1-46.

Bentley, Delia, 2004. « Definiteness Effects: Evidence from Sardinian ». *Transactions of the Philological Society* 102(1) : 57-101.

Berlin, Brent et Paul Kay, 1969. *Basic Color Terms: Their Universality and Evolution*. Berkeley : University of California Press.

Boerger, Brenda H, 1984. *Proto-Romanes phonology*. Thèse de doctorat, Austin : The University of Texas. SIL Language and Culture Archives.

Boretzky, Norbert, 1986. « Zur Sprache der Gurbet von Priština (Jugoslawien) ». *Gießener Hefte für Tsiganologie* 3(1–4) : 195-217.

———, 1993. *Bugurdzi: Deskriptiver und historischer Abriss eines Romani-Dialekts*. Wiesbaden : Harrassowitz in Kommission.

———, 1994. *Romani: Grammatik des Kalderaš-Dialekts*. Berlin : Harrassowitz Verlag.

———, 1996. « The “New” Infinitive in Romani ». *Journal of the Gypsy Lore Society* 6(1) : 1-51.

———, 1998. « Areal and Insular Dialects and the Case of Romani ». *Grazer Linguistische Studien* (50) : 1-27.

———, 2000a. « South Balkan II as a Romani Dialect Branch: Bugurdži, Drindari and Kalajdži ». *Romani Studies* 10(2) : 105-183.

———, 2000b. « The Definite Article in Romani Dialects ». *Grammatical Relations in Romani: The Noun Phrase*, édité par Yaron Matras et Viktor Elšík. 31-64. Amsterdam & Philadelphia : John Benjamins Publishing Company.

———, 2003. *Die Vlach-Dialekte des Romani : Strukturen, Sprachgeschichte, Verwandtschaftsverhältnisse, Dialektkarten*. Wiesbaden : Harrassowitz Verlag.

———, 2007. « The Differentiation of the Romani Dialects ». *Sprachtypologie und Universalienforschung* 60(4) : 314-336.

Boretzky, Norbert et Birgit Igla, 2004. *Kommentierter Dialektatlas des Romani*. Wiesbaden : Harrassowitz Verlag.

Börjars, Kersti, 1994. « Swedish Double Determination in a European Typological Perspective ». *Nordic Journal of Linguistics* 17 : 219-252.

Borsley, Robert D. et Ian Roberts, 1996. « Introduction ». *The Syntax of the Celtic Languages*, édité par Robert D. Borsley et Ian Roberts. 9-74. Cambridge : Cambridge University Press.

Bouchard, Denis, 1998. « The Distribution and Interpretation of Adjectives in French: A Consequence of Bare Phrase Structure ». *Probus* 10(2) : 139-183.

Bourcart, Jacques. 1922. *Les Confins Albanais administrés par la France (1916-1920) : contribution à la géographie et à la géologie de l'Albanie moyenne*. Thèse, Paris : Université de Paris.

Breu, Walter, 1996. « Überlegungen zu einer Klassifizierung des grammatischen Wandels im Sprachkontakt (am Beispiel slavischer Kontaktfälle) ». *Sprachtypologie und Universalienforschung* 49(1) : 21-38.

Büring, Daniel, 2009. « Towards a Typology of Focus Realization ». *Information Structure*, édité par Malte Zimmerman et Caroline Féry. 177-205. Oxford : Oxford University Press.

Cabredo Hoffherr, Patricia et Ora Matushansky, éd. 2010. *Adjectives: Formal analyses in Syntax and Semantics*. Amsterdam & Philadelphia : John Benjamins Publishing Company.

Cabredo Hofherr, Patricia et Anne Zribi-Hertz, éd. 2014. *Crosslinguistic Studies on Noun Phrase Structure and Reference*. Leiden : Brill.

Campos, Héctor, 2005. « Noun Modification, Pseudo-Articles, and Last Resort Operations in Arvantovlaxika and in Romanian ». *Lingua* (115) : 311-347.

———, 2009. « Some Notes on Adjectival Articles in Albanian ». *Lingua* (119) : 1009-1034.

Campos, Héctor et Melita Stavrou, 2004. « Polydefinite Constructions in Modern Greek and in Aromanian ». *Balkan Syntax and Semantics*, édité par Olga Mišeska Tomić. 137-173. Amsterdam & Philadelphia : John Benjamins Publishing Company.

Cech, Petra, 2006. *Dolenjska Romani: The Dialect of the Dolenjski Roma in Novo Mesto and Bela Krajina, Slovenia*. München : Lincom.

Cech, Petra et Mozes F. Heinschink, 1999. *Sepečides-Romani*. *Balkanologische Veröffentlichungen, 34.* Wiesbaden : Harrassowitz Verlag.

———, 2001. « A Dialect with Seven Names ». *Romani Studies Fifth Series* (11) : 137-184.

Çelebi, Evliya, 1630. سیاحت نامه, / *Seyāḥat-Nāme* / *Livre des Voyages*.

Chafe, Wallace L., 1976. « Givenness, Contrastiveness, Definiteness, Subjects, Topics and Point of View ». *Subject and Topic*, édité par Charles N. Li. 25-55. New York : Academic Press.

Chambers, John Kenneth, 2013. « Patterns of Variation Including Change ». *The Handbook of Language Variation and Change*, 2ème édition, édité par John Kenneth Chambers et Nathalie Schilling. 297-323. Oxford : Blackwell Publishing.

Charolles, Michel, 2002. *La référence et les expressions référentielles en français*. Paris : Ophrys.

Chatzikyriakidis, Stergios. 2014. « Polydefinites in Modern Greek: Where Restrictive Interpretation Fails ». Ms. Laboratoire d'Informatique, de Robotique et de Microélectronique de Montpellier.

———, 2015. « Polydefinites in Modern Greek: The Missink LINK ». *Proceedings of the 11th International Conference on Greek Linguistics* : 1-22.

———, 2016. « Polydefinites in Greek: When Restrictive Interpretation fails ». Présentation au Linguistics Seminar of the English Department of Humboldt University, Berlin, juin.

Ciccozzi, Massimo, Anna-Rita Ciccaglione, Alessandra Lo Presti et al., 2011. « Reconstruction of the Evolutionary Dynamics of the Hepatitis C Virus 1b Epidemic in Turkey ». *Infection, Genetics and Evolution* 11(5) : 863-868.

Cinque, Guglielmo, 2005. « Deriving Greenberg's Universal 20 and its Exceptions ». *Linguistic Inquiry* 36(3) : 315-332.

———, 2010. *The Syntax of Adjectives: A Comparative Study*. Cambridge, Massachusetts : The MIT Press.

———, 2014. « The Semantic Classification of Adjectives. A View from Syntax ». *Studies in Chinese Linguistics* 35(1) : 3-32.

Cioabă, Luminița Mihai, 2006a. *Lacrimi Rome - Romane Asva*. București : Ro Media.

———, 2006b. *Romane Iasfa - Gypsy Tears.* Documentary film, Interviews. Social-Cultural Foundation of Roma "Ion Cioaba."

Conseil de l'Europe, 2012. *Glossaire terminologique raisonné du conseil de l'Europe sur les questions roms.* https://www.ovr-suisse.ch/sites/default/files/glossaire_roms_fr_version_18_may_2012.pdf

Constantinescu, Barbu, 2016. *Cântece Țigănești - Romané Ghilea - Gypsy Songs*, édité par Julieta Rotaru. București : Editura muzeul literaturii Române.

Coppolani, Marie-Laure, 2018. *Le changement lexical en nez-percé*. Thèse de doctorat, Paris : Institut National des Langues et Civilisations Orientales.

Corbett, Greville G., 2006. *Agreement*. Cambridge : Cambridge University Press.

Cornilescu, Alexandra et Alexandru Nicolae, 2012. « Nominal Ellipsis as Definiteness and Anaphoricity: The Case of Romanian ». *Lingua* 122 : 1070-1111.

Courthiade, Marcel, 1985. « Tatomir Vukanović - Romi (Cigani) u Jugoslaviji ». *Études Tsiganes* (3) : 52.

———, 1986a. « Histoire de compléter Sandfeld : le rromani comme langue balkanique ».

———, 1986b. « The Balkanicity of the Rromani Language ». Présentation au Congrès de l'Association internationale des études du Sud-Est européen.

———, 1988a. « Les derniers vestiges du parler slave de Bobosëica et de Drenovene (Albanie) ». *Revue des études slaves* 60 (1) : 139-57.

———, 1988b. « Rajko Đurić - Seobe Roma ». *Études Tsiganes* (1) : 51-52.

———, 1995. « Between Conviviality and Antagonism: The Ambiguous Position of the Romanies in Albania ». *Patrin* (3).

———, 2004. *Fjalor rromisht-shqip: i të folmeve Rrome në përdorim në Shqipëri: meçkarisht, kabuxhisht, rupanisht e shkodranisht*. Tiranë : Rromani Baxt.

———, 2007. *La littérature des Rroma, Sintés et Kalés : compendium à l'usage des étudiants de l'Inalco, section Langue et Civilisation Rromani et filière Français Langue Étrangère*. Paris : Inalco.

———, 2009a. *Introduction à la langue rromani - La langue rromani, présentation générale*. Paris : Inalco - D.A.S. - C.E.B. - F.L.E.

———, 2009b. *Morri angluni rromane ćhibăqi evroputni lavustik / első rromani nyelvű európai szótáram: cigány, Mmagyar, angol, francia, spanyol, német, ukrán, román, horvát, szlovák, görög*, édité par Melinda Rézműves, Lucia Presber et Geoff Husic. Budapest : Fővárosi Onkormányzat Cigány Ház - Romano Kher.

———, 2013. « Le rromani et les autres langues en usage parmi les Rroms, Manouches et Gitans ». *Pour une histoire sociale des langues de France*, édité par Georg Kremnitz. 567-580. Rennes : Presses Universitaires de Rennes.

———, 2014a. « Les premières approches linguistiques du rromani (1500-1800) : entre présupposés, dégoût, ambitions et objectivité de méthode ». *Dossiers d'HEL, SHESL. Linguistiques d'intervention. Des usages socio-politiques des savoirs sur le langage et les langues* : 1-10.

———, 2014b. « Standardisation du rromani : la foire aux malentendus ». Ms. Institut National des Langues et Civilisations Orientales, Paris.

Creissels, Denis, 2006. *Syntaxe générale, une introduction typologique. Tome 2 : la phrase*. Paris : Hermes Science Publications.

Dahl, Östen, 2004. « Definite Articles in Scandinavian: Competing Grammaticalization Processes in Standard and Non-Standard Varieties ». *Dialectology Meets Typology*, édité par Bernd Kortmann. 147-80. *Trends in Linguistics*, 153. Berlin & New York : Mouton de Gruyter.

De Mulder, Walter et Anne Carlier, 2011. « Grammaticalization of Definite Articles ». *The Oxford Handbook of Grammaticalization*, édité par Bernd Heine et Heiko Narog. 522-534. Oxford : Oxford University Press.

De Soto, Hermine, Sabine Beddies et Ilir Gedeshi, 2005. *Roma and Egyptians in Albania: From Social Exclusion to Social Inclusion*. Washington D.C. : The International Bank for Reconstruction and Development / The World Bank. http://siteresources.worldbank.org/EXTROMA/Resources/2005_Albania_WB_Report_Social_Exclusion.pdf

Deman, Mag. Katharina. 2005. *Untersuchung zur Grammatik der Romani – Varietät der Kumpanja in Bogotá/Kolumbien*. Thèse de doctorat, Graz : Karl-Franzens-Universität.

Denison, Norman, 1969. « Sociolinguistic Aspects of Plurilingualism ». *Giornate internazionali di sociolinguistica*. 255-278. Roma : Istituto L. Sturzo.

———, 1971. « Some Observations on Language Variety and Plurilingualism ». *Social Anthropology and Language*, édité par Edwin Ardener. 157-183. London : Tavistock Publications.

Despić, Miloje, 2013. « Binding and the Structure of NP in Serbo-Croatian ». *Linguistic Inquiry* 44(2) : 239-270.

Dik, Simon C., 1997. *The Theory of Functional Grammar*. Berlin : Walter de Gruyter.

Dobrushi, Andi, 2013. *Roma Decade and the Situation of Roma Community in Albania*. Tiranë : The Open Society Foundations in Albania.

Draganova, Desislava, 2005. « Turkish Verbs in Bulgarian Romani ». *General and Applied Romani Linguistics. Proceedings from the 6th International Conference on Romani Linguistics*, édité par Barbara Schrammel, Dieter W. Halwachs et Gerd Ambrosch. 90-98. München : Lincom Europa.

Durand, Tom, 2016. *L'intransitivité scindée dans les langues arawak*. Thèse de doctorat, Paris : Institut National des Langues et Civilisations Orientales.

Dryer, Matthew S., 2013. « Definite Articles ». *The World Atlas of Language Structures* Online, édité par Matthew S. Dryer et Martin Haspelmath. Leipzig : Max Planck Institute for Evolutionary Anthropology. http://wals.info/chapter/37

Dubois, Jean, Mathée Giacomo, Louis Guespin et al., 2012. *Le dictionnaire de linguistique et des sciences du langage*. Grands dictionnaires Larousse. Paris : Larousse.

Durham, Mary Edith, 1905. *The Burden of the Balkans*. London : Edward Arnold.

Đurić, Rajko, 1983. « Romski jezik. Gurbetski dijalekat ». *Zagonetki motivi i jezik Roma*, édité par Rajko Đurić. 43-59. Beograd : Bagdala.

Eckert, Penelope, 2000. *Linguistic Variation as Social Practice.* Oxford : Blackwell Publishing.

Elšík, Viktor, 2000. « Dialect Variation in Romani Personal Pronouns ». *Grammatical Relations in Romani: The Noun Phrase*, édité par Viktor Elšík et Yaron Matras. 65-94. Amsterdam & Philadelphia : John Benjamins Publishing Company.

Erteschik-Shir, Nomi, 2014. « Information Structure and (in)Definiteness ». *Crosslinguistic Studies on Noun Phrase Structure and Reference*, édité par Patricia Cabredo Hofherr et Anne Zribi-Hertz. 23-51. Leiden : Brill.

Etxeberria, Urtzi et Anastasia Giannakidou, 2014. « D-Heads, Domain Restriction and Variation: From Greek and Basque to Salish ». *Nominal Constructions in Slavic and Beyond*, édité par Lilia Schürcks, Anastasia Giannakidou, Urtzi Etxeberria et Peter Kosta. 413-440. Berlin & Boston : Mouton de Gruyter.

European Roma Rights Center, 2010. *Basic Facts on Roma - Albania.* http://www.errc.org/article/basic-facts-on-roma--albania/3622

Evert, Stefan, 2006. « How Random is a Corpus ? The Library Metaphor ». *Zeitschrift Für Anglistik Und Amerikanistik* 54(2) : 177-190.

Fabri, Ray, 2001. « Definiteness Marking and the Structure of the NP in Maltese ». *Verbum* 2 : 153-172.

Fagard, Benjamin. 2015. *Pour une romanistique moderne : la microtypologie*. Habilitation à diriger des recherches, Tours : Université François-Rabelais.

Ferková, Ilona, 1996. *Corde chave - ukradené deti*. Brno : Romano jekhetaniben pre Morava.

Feuillet, Jack, 2012. *Linguistique comparée des langues balkaniques*. Paris : Institut d'Études Slaves.

Firbas, Jan, 1964. « On Defining the Theme in Functional Sentence ». *Travaux Linguistiques de Prague* 1 : 267-280.

Forsgren, Mats, 1978. *La place de l'adjectif épithète en français contemporain, étude quantitative et sémantique*. Stockholm : Almqvist & Wilksell.

———, 1997. « Un classique revisité : la place de l'adjectif épithète ». *Études de linguistique française, médiévale et générale offertes à Robert Martin à l'occasion de ses 60 ans*, édité par Georges Kleiber et Martin Riegel. 115-126. Louvain-la-Neuve : Duculot.

Franco, Ludovico, Rita Manzini et Leonardo Maria Savoia, 2013. « Some Call It Article, Some Ezafe: 'Linkers' as 'Agreement' ». http://ling.auf.net/lingbuzz/001888

Foulkes, Paul et Gerard J. Docherty, 1999. « Urban Voices – Overview ». *Urban Voices*, édité par Paul Foulkes et Gerard J. Docherty. 1-24. London : Edward Arnold.

Foulon, Jordanka, 1997. « Les articles définis en macédonien ». *Cahiers Balkaniques* 23 : 7-30.

Friedman, Eben, 2015. *Decade of Roma Inclusion: Progress Report*. Tiranë : Programme des Nations Unies pour le développement.

Friedman, Victor A., 1993. « Dialect Variation and Questions of Standardization in Macedonia: Macedonian, Albanian and Romani ». *Zbornik za filologiju i lingvistiku* 36. (2) : 7-35.

———, 1999. « The Bulgarian Indefinite Article Revisited in its Balkan Context ». *Bulgarian Studies at the Dawn of the 21st Century: A Bulgarian-American Perspective. (Sixth Joint Meeting of Bulgarian and North American Scholars. Blagoevgrad, Bulgaria)*, édité par Anisava Miltenova : 105-114.

———, 2000a. « Confirmative/Nonconfirmative in Balkan Slavic, Balkan Romance, and Albanian, with Additional Observations on Turkish, Romani, Georgian, and Lak ». *Evidentials. Turkic, Iranian and Neighbouring Languages*, édité par Lars Johanson et Bo Utas. Berlin & Boston : Mouton de Gruyter.

———, 2000b. « Observations on the Use of Jedan as a Marker of Indefiniteness in Serbian/Croatian ». *In a Foreign Harbor: Essays in Honor of Vasa D. Mihajlović*, édité par Radmila Gorup et Bogdan Rakić. 195-204. Bloomington : Slavica.

———, 2000c. « Romani in the Balkan Linguistic League ». *Valkanikē Glōssologia: Sygkhronia kai Diakhronia/Balkanlinguistik: Synchronie und Diachronie*, édité par Christos Tzitzilis et Charalambos Symeonidēs.95-105. Thessaloniki : University of Thessaloniki.

———, 2001. « The Romani Indefinite Article in its Historical and Areal Context ». *Was ich noch sagen wollte... A multilingual festschrift for Norbert Boretzky on the occasion of his 65th birthday*, édité par Birgit Igla et Thomas Stolz. 287-301. *Sprachtypologie und Universalienforschung*, Beihefte, Studia Typologica, 2. Berlin : Akademie Verlag.

———, 2003. « Turkish in Macedonia and Beyond: Studies in Contact, Typology and Other Phenomena in the Balkans and the Caucasus ». *Turcologica*, 52. Wiesbaden : Harrassowitz Verlag.

———, 2003a. « “One” as an Indefinite Marker in Balkan and Non-Balkan Slavic ». *Prilozi: oddelenie za lingvistika i literaturna nauka – Makedonska Akademija na Naukite i Umetnostite* 28(1) : 109-151

———, 2003b. « The Indefiniteness of “one” in its Macedonian and Balkan Context ». *Zbornik referati od IV makedonsko-severnoamerikanska slavistička konferencija za makedonistika*, édité par Vera Stoječevska-Antik. 25-33. Skopje : University of Skopje.

———, 2006. « West Rumelian Turkish in Macedonia and Adjacent Areas ». *Turkic Language Contacts*, édité par Hendrik Boeschoten et Lars Johanson. 27-45. Wiesbaden : Harrassowitz Verlag.

———, 2017. « Seven Varieties of Arli: Skopje as a Center of Convergence and Divergence of Romani Dialects ». *Romani Studies* 27(1) : 29-45.

Friedman, Victor A. et Robert Dankoff, 1991. « The Earliest Text in Balkan (Rumelian) Romani : A Passage from Evliya Çelebi’s Seyahat-Nameh ». *Journal of the Gypsy Lore Society* (Fifth Series) 1(1) : 1-20.

Gaatone, David, 2016. « Les facettes obscures de l'adjectif en français ». *Sur les traces de l'adjectif*, édité par Spitzl-Dupic, Friederike, Michaël Grégoire, Lidia Lebas-Fraczak et Richard Ryan. 15-25. Clermont-Ferrand : Presses Universitaires Blaise Pascal.

Gaiffe, Bertrand, Anne Reboul et Laurent Romary, 1997. « Les SN définis : anaphore, anaphore associative et cohérence ». *Relations anaphoriques et (in)cohérence*, édité par Walter De Mulder, Liliane Tasmowski-De-Ryck et Carl Vetters. 69-97. Amsterdam : Rodopi.

Gardiner, Alan H., 1932. *The Theory of Speech and Language*. Oxford : Clarendon Press.

———, 1989. *Langage et acte de langage, aux sources de la pragmatique*, traduit par Catherine Douay. Lille : Presses Universitaires de Lille.

Garo, Morgan, 2002. « La langue rromani au cœur du processus d'affirmation de la nation rrom ». *Hérodote* 2(105) : 154-165.

Gëdeshi, Ilir et Juna Miluka, 2012. *A Needs Assessment Study on Roma and Egyptian Communities in Albania*, traduit par Majlinda Nishku. Tirana : UNDP (United Nations Development Programme) - Albania. http://www.al.undp.org/content/dam/albania/docs/Roma%20Needs%20Assessment.pdf

Giannakidou, Anastasia et Melita Stavrou, 1999. « Nominalization and Ellipsis in the Greek DP ». *The Linguistic Review* 16 : 295-331.

Gilliat-Smith, Bernard J., 1915. « A Report on the Gypsy Tribes of North East Bulgaria ». *Journal of the Gypsy Lore Society* 9 : 1-55.

Giusti, Giuliana, 1995. « A Unified Structural Representation of (Abstract) Case and Articles: Evidence from Germanic ». *Studies in Comparative Germanic Syntax*, édité par Hubert Haider, Susan Olsen et Sten Vikner. 77-93. Dordrecht : Kluwer.

Givón, Talmy, 1984. *Syntax: A Functional-Typological Introduction, vol.1*. Amsterdam & Philadelphia : John Benjamins Publishing Company.

———, 1990. *Syntax: A Functional-Typological Introduction., vol.2*. Amsterdam & Philadelphia : John Benjamins Publishing Company.

Gjerde, Lars et Knut Kristiansen, 1994. *The Orange of Love and Other Stories: The Rom-Gypsy Language in Norway*. Oslo : Scandinavian University Press.

Gjerdman, Olof et Erik Ljungberg, 1963. « The Language of the Swedish Coppersmith Gipsy Johan Dimitri Taikon: Grammar, Texts, Vocabulary and English Word-Index ». *Acta Academiae Regiae Gustavi Adolphi*, 40. Uppsala : Lundequistska bokh.

Goes, Jan, 1999. *L'adjectif : entre nom et verbe*. Louvain-la-Neuve : De Boeck-Duculot.

Götze, Michael, Thomas Weskott, Cornelia Endriss et al., 2007. « Information Structure ». *Interdisciplinary Studies on Information Structure* (07) : 147-187.

Greenberg, Joseph H., 1963. « Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements ». *Universals of Grammar*, édité par Joseph H. Greenberg. 73-113. Cambridge, Massachusetts : MIT Press.

———, 1978. « How Does a Language Acquire Gender Markers? ». *Universals of Human Language: Word Structure*, édité par Joseph H. Greenberg, Charles A. Ferguson et Edith A. Moravcsik. 47-82. Stanford : Stanford University Press.

Gresham, David, Bharti Morar, Peter A. Underhill et al., 2001. « Origins and Divergence of the Roma (Gypsies) ». *American Journal of Human Genetics* 69 : 1314-1331.

Gundel, Jeanette K., 1999. « On Different Kinds of Focus ». *Focus: Linguistic, Cognitive and Computational Perspectives*, édité par Peter Bosch et Rob van der Sandt. 293-305. Cambridge : Cambridge University Press.

Gundel, Jeanette K. et Thorstein Fretheim, 2006. « Topic and Focus ». *The Handbook of Pragmatics*, édité par Laurence R. Horn et Gregory Ward. 175-296. Oxford : Blackwell Publishing.

Gurbetovski [Courthiade], Medo [Marcel], 2010. *Rromani de poche*. Chennevières-sur-Marne : Assimil.

Hackaj, Krisela, 2014. *Needs Assessment of Roma and Egyptian Individuals on Employment and Entrepreneurship in the Regions of Korca, Berat and Vlora*. Tiranë : Programme des Nations Unies pour le développement.

Halliday, Michael, 1967. *Intonation and Grammar in British English*. Den Haag : Mouton de Gruyter.

———, 1985. *An Introduction to Functional Grammar*, 3ème édition, éditée par Christian M. Matthiessen. London : Hodder Arnold.

Halwachs, Dieter W. 1998. *Amaro vakeripe roman hi – Unsere Sprache ist Roman: Texte, Glossar und Grammatik der burgenländischen Romani-Variante*. Klagenfurt : Drava.

Halwachs, Dieter W, Gerd Ambrosch, Katharina Deman et Ursula Glaeser. 2002. *Burgenland-Romani*. München : LINCOM Europa.

Hasluck, Margaret, 1938a. « The Gypsies of Albania - I ». *Journal of the Gypsy Lore Society* (Third Series) 17(2) : 49-61.

———, 1938b. « The Gypsies of Albania - III ». *Journal of the Gypsy Lore Society* (Third Series) 17(4) : 110-122.

Haßler, Gerda, 2016. « Analyse sémantico-contrastive du rôle des adjectifs qualificatifs ». *Sur les traces de l'adjectif*, édité par Spitzl-Dupic, Friederike, Grégoire, Michaël, Lidia Lebas-Fraczak et Ryan, Richard. 27-44. Clermont-Ferrand : Presses Universitaires Blaise Pascal.

Hawkins, John A., 1978. *Definiteness and Indefiniteness: A Study in Reference and Grammaticality Prediction*. London : Croom Helm.

Hays, Jo N., 2005. *Epidemics and Pandemics - Their Impacts on Human History*. Santa Barbara, California : ABC-CLIO.

Heine, Bernd, 2008. « Contact-Induced Word Order Change without Word Order Change ». *Language Contact and Contact Languages*, édité par Peter Siemund et Noemi Kintana. 33-60. Amsterdam & Philadelphia : John Benjamins Publishing Company.

Heine, Bernd et Tania Kuteva, 2005. *Language Contact and Grammatical Change*. Cambridge : Cambridge University Press.

Heinschink, Mozes F., 1978. « La langue tsigane parlée en Autriche et en Yougoslavie. *Études Tsiganes* (1) : 8-20.

Herin, Bruno, 2012. « The Domari Language of Aleppo (Syria) ». *Linguistic Discovery* 10(2) : 1-52.

———, 2014. « Le marquage différentiel en domari ». Présentation aux Séances scientifiques de la Société Linguistique de Paris, EPHE, Paris, 21 juin.

Hohmann, Joachim Stephan, 1990. *Verfolgte ohne Heimat: Geschichte der Zigeuner in Deutschland.* Frankfurt am Main & New York : Peter Lang.

Hoxha, Enver, 1984. « Les méprisés sont devenus heureux », traduit par Marcel Courthiade. *Études Tsiganes* 1 : 26-31.

Igla, Birgit, 1996. *Das Romani von Ajia Varvara: deskriptive und historisch-vergleichende Darstellung eines Zigeunerdialekts*. Wiesbaden : Otto Harrassowitz Verlag.

Ioannidou, Alexandra et Marcel Den Dikken, 2009. « P-Drop, D-Drop, D-Spread ». *MIT Working Papers in Linguistics* (57) : 393-408.

Jacobs, Neil G., 2005. *Yiddish: A Linguistic Introduction*. Cambridge : Cambridge University Press.

Jahan, Selim et Eva Jespersen, 2015. *Rapport sur le développement humain*. New York : Programme des Nations Unies pour le développement. http://hdr.undp.org/sites/default/files/fr_hdr_2015_1021_web.pdf

Jick, Todd D., 1979. « Mixing Qualitative and Quantitative Methods: Triangulation in Action ». *Administrative Science Quarterly* 24(4) : 602-611.

Joseph, Brian D., 2001. « Is Balkan Comparative Syntax Possible? ». *Comparative Syntax of Balkan Languages*, édité par Maria-Luisa Rivero et Angela Ralli. 17-43. Oxford & New York : Oxford University Press.

Kamp, Hans, 1975. « Two Theories about Adjectives ». *Formal Semantics of Natural Language*, édité par Edward Keenan.123-155. Cambridge : Cambridge University Press.

Karatsareas, Petros, 2011. *A Study of Cappadocian Greek Nominal Morphology from a Diachronic and Dialectological Perspective*. Thèse de doctorat, Cambridge : University of Cambridge.

———, 2013. « Understanding Diachronic Change in Cappadocian Greek: The Dialectological Perspective ». *Journal of Historical Linguistics* 3 : 192-229.

Karatsareas, Petros, et Marika Lekakou. 2014. « Determiner Spreading in Cappadocian Greek ». Présentation à la 6th International Conference on Modern Greek Dialects and Linguistic Theory, University of Patras, Patras, 27 septembre.

Kariaeva, Natalia. 2004. *Agreement and Locality*. Thèse de doctorat, New Jersey : Rutgers University

Koinova, Maria, 2000. *Roma of Albania. Minorities in Southeast Europe.* Athènes : CEDIME-SE (Center for Documentation and Information on Minorities in Europe-Southeast Europe), Greek Helsinki Monitor. www.greekhelsinki.gr/pdf/cedime-se-albania-roma.doc

Kehrein, Roland, 2002. *Prosodie und Emotionen*. Tübingen : Mouton de Gruyter.

Kolliakou, Dimitra, 1998. « Linkhood and Multiple Definite Marking ». *Proceedings of FHCG'98, (Joint Conference on Formal Grammar, Head-Driven Phrase Structure Grammar, and Categorial Grammar)*, édité par Gosse Bouma, Geert-Jan M. Kruijff et Richard Oehrle. 14-22. Saarbrücken.

———, 2003. *Nominal Constructions in Modern Greek: Implications for the Architecture of Grammar*. Stanford, California : CSLI Publications (Center for the Study of Language and Information).

———, 2004. « Monadic Definites and Polydefinites: Their Form, Meaning and Use ». *Journal of Linguistics* 40 : 263-333.

Kolsti, John, 1991. « Albanian Gypsies, The Silent Survivors ». *The Gypsies of Eastern Europe*, édité par David Crowe et John Kolsti. Armonk, New York : M.E. Sharpe.

Koptjevskaja-Tamm, Maria, 2000. « Romani Genitives in Cross-Linguistic Perspective ». *Grammatical Relations in Romani: The Noun Phras*e, édité par Yaron Matras et Viktor Elšík. 123-150. Amsterdam Studies in the Theory and History of Linguistic Science. Amsterdam & Philadelphia : John Benjamins Publishing Company.

Kyriakaki, Maria, 2010. « What Greek DETs Do: The Restrictive DP ». *Actes du congrès annuel de l'Association canadienne de linguistique 2010 / Proceedings of the 2010 Annual Conference of the Canadian Linguistic Association*. Montréal : Melinda Heijl.

———, 2011. *DETs in the Functional Syntax of Greek Nominals*. Thèse de doctorat, Toronto : University of Toronto. https://tspace.library.utoronto.ca/bitstream/1807/31812/1/Kyriakaki_Maria_201112_PhD_thesis.pdf

———, 2014a. « The Three Degrees of Definiteness ». *Proceedings of the Fortieth Annual Meeting of the Berkeley Linguistics Society*, édité par Herman Leung, Zachary O'Hagan, Sarah Bakst et al. 253-269. Berkeley.

———, 2014b. « Definiteness and Polydefinites ». Présentation au stage doctoral du SeDyL, CNRS, Villejuif, 4 juin. https://stagesedyl.hypotheses.org/programme-et-resumes

Labov, William, 1972. *Sociolinguistic Patterns, 4.* Philadelphia : University of Pennsylvania Press.

———, 1994. *Principles of Linguistic Change. Volume I: Internal Factors.* Oxford : Oxford University Press.

Lacito, UMR 7107, 2001. *Collection Pangloss.* La Collection Pangloss. http://lacito.vjf.cnrs.fr/pangloss/

Laenzlinger, Christopher, 2000. « French Adjective Ordering: Perspectives on DP-Internal Movement Types ». *Generative Grammar in Geneva* 1 : 55-104.

———, 2005. « French Adjective Ordering: Perspectives on DP-Internal Movement Types ». *Lingua* 115(5) : 645-689.

Lambrecht, Knud, 1994. *Information Structure and Sentence Form: Topic, Focus, and the Mental Representations of Discourse Referents*. Cambridge : Cambridge University Press.

Laury, Ritva, 1997. *Demonstratives in Interaction: The Emergence of a Definite Article in Finnish.* Amsterdam & Philadelphia : John Benjamins Publishing Company.

Lebas-Fraczak, Lidia, 2009. « Description "communicative" des déterminants français en vue de la didactisation ». *Les Cahiers de l'Acedle, didactique des langues et linguistique* 6(2) : 1-19.

———, 2015. *Pour une grammaire interlocutive du français*. Habilitation à diriger des recherches, Clermont-Ferrand : Université Blaise Pascal.

———, 2017. « Interaction entre l'article et l'adjectif dans le marquage de la focalisation au sein des propositions attributives ». *Points de vue : mélanges offerts à Henning Nølke à l'occasion de sa retraite*, édité par Merete Birkelund. 75-89. Aarhus : IKK, Aarhus Universitet.

Leggio, Daniele Viktor, 2011. « The Dialect of the Mitrovica Roma ». *Romani Studies* 5(21), 57-114.

Lekakou, Marika et Petros Karatsareas, 2016. « Marking Definiteness Multiply: Evidence from Two Varieties of Greek ». *Studies in Greek Linguistics* (36) : 189-204.

Lekakou, Marika et Kriszta Szendrői, 2007. « Eliding the Noun in Close Apposition, or Greek Polydefinites Revisited ». *UCL Working Papers in Linguistics* 19 : 129-154.

———, 2009. « Close Apposition with and without Noun Ellipsis: An Analysis of Greek Polydefinites ». *Studies in Greek Linguistics* (29) : 151-166.

———, 2012. « Polydefinites in Greek: Ellipsis, Close Apposition and Expletive Determiners ». *Journal of Linguistics* (48) : 107-149.

———, 2014. « When Determiners Abound: Implications for the Encoding of Definiteness ». *Crosslinguistic Studies on Noun Phrase Structure and Reference*, édité par Patricia Cabredo Hoffherr et Anne Zribi-Hertz. 212-238. Leiden : Brill.

Leu, Thomas, 2009. « From Greek to Germanic: Poly-(*in)-definiteness and Weak/Strong Adjectival Inflection ». *Merging Features: Computation, Interpretation, and Acquisition*, édité par José M. Brucart, Anna Gavarró et Jaume Solà. 293-309. Oxford : Oxford University Press.

Löbner, Sebastian, 1985. « Definites ». *Journal of Semantics* 4 : 279-326.

Lyons, Christopher, 1985. « A Possessive Parameter ». *Sheffield Working Papers in Language and Linguistics* 2 : 98-104.

———, 1986. « The Syntax of English Genitive Constructions ». *Journal of Linguistics* 22 : 123-43.

———, 1999. *Definiteness*. Cambridge : Cambridge University Press.

Mann, Stuart E., 1933a. « Albanian Romani - I ». *Journal of the Gypsy Lore Society*, Third Series 12(1) : 1-32.

———, 1933b. « Albanian Romani - II ». *Journal of the Gypsy Lore Society*, Third Series 12(3) : 147-152.

———, 1934a. « Albanian Romani - III ». *Journal of the Gypsy Lore Society*, Third Series 13(1) : 37-45.

———, 1934b. « Albanian Romani - IV ». *Journal of the Gypsy Lore Society*, Third Series 13(3) : 153-158.

———, 1935. « South Albanian Romani - I ». *Journal of the Gypsy Lore Society*, Third Series 14(4) : 174-184.

———, 1937. « South Albanian Romani - II ». South Albanian Romani Texts. *Journal of the Gypsy Lore Society*, Third Series 16(4) : 186-195.

———, 1941. « Albanian Romani - V. North Albanian Romani Texts ». *Journal of the Gypsy Lore Society*, Third Series 20(1–2) : 85-87.

———, 1990. *Stuart Manns Wörterbuch Des Albanischen Romanes*, édité par Marcel Courthiade. Tsiganologische Studien, 1. Gießen : Justus-Liebig-Universität Gießen, Institut für Soziologie.

Manolessou, Io, 2000. *Greek Noun Phrase Structure: A Study in Syntactic Evolution*. Thèse de doctorat, Cambridge : University of Cambridge.

Marchis, Mihalea et Artemis Alexiadou, 2009. « On the Distribution of Adjectives in Romanian: The *Cel* Construction ». *Romance Languages and Linguistic Theory*, édité par Enoch O. Aboh, Elisabeth van der Linden, Josep Quer et Petra Sleeman. 161-178. Amsterdam & Philadelphia : John Benjamins Publishing Company.

Marinis, Theodoros, 2003. *The Acquisition of the DP in Modern Greek*. Amsterdam & Philadelphia : John Benjamins Publishing Company.

Marit, Julien, 2005. *Nominal Phrases from a Scandinavian Perspective*. Amsterdam & Philadelphia : John Benjamins Publishing Company.

Marushiakova, Elena et Vesselin Popov, 2016. « Identity and Language of the Roma (Gypsies) in Central and Eastern Europe ». *The Palgrave Handbook of Slavic Languages, Identities and Borders*, édité par Tomasz Kamusella, Motoki Nomachi et Catherine Gibson. 26-54. Basingstoke : Palgrave Macmillan.

Marušič, Franc et Rok Žaucer, 2006. « The Definite Article in Colloquial Slovenian and an AP Related DP Position ». *Formal Approaches to Slavic Linguistics 14*, édité par James Lavine, Steven Franks, Mila Tasseva-Kurktchieva et Hana Filip. 189-204. Ann Arbor : Michigan Slavic Publications.

———, 2008. « On the Adjectival Definite Article in Slovenian ». *Pismo* 5(1) : 102-124.

———, 2014. « A Definite Article in the AP - Evidence from Colloquial Slovenian ». *Nominal Constructions in Slavic and Beyond*, édité par Lilia Schürcks, Anastasia Giannakidou, Urtzi Etxeberria et Peter Kosta. 183-208. Berlin & Boston : Mouton de Gruyter.

Masica, Colin P., 1991. *The Indo-Aryan Languages*. Cambridge : Cambridge University Press.

Master, Peter, 1997. « The English Article System: Acquisition, Function, and Pedagogy ». *System* 25(2) : 215-232.

Matras, Yaron. 1994. *Untersuchungen zu Grammatik und Diskurs des Romanes: Dialekt der Kelderasa/Lovara*. Wiesbaden : Harrassowitz.

———, 1996. « Prozedurale Fusion: Grammatische Interferenzschichten im Romanes ». *Sprachtypologie und Universalienforschung* 49(1) : 60-78.

———, 1997. « The Typology of Case and Case Layer Distribution in Romani ». *The Typology and Dialectology of Romani*, édité par Yaron Matras, Peter Bakker et Hristo Kyuchukov. 61-93. Amsterdam & Philadelphia : John Benjamins Publishing Company.

———, 1999a. « The Speech of the Polska Roma: Some Highlighted Features and Their Implications for Romani Dialectology ». *Journal of the Gypsy Lore Society* 5(9) : 1-28.

———, 1999b. « The State of Present-Day Domari in Jerusalem ». *Mediterranean Language Review* 11 : 1-58.

———, 2000a. « Migrations and “Replacive Convergence” as Sources of Diversity in the Dialects of Romani ». *Dialect and Migration in a Changing Europe*, édité par Klaus Mattheier. 173-194. Frankfurt am Main & New York : Peter Lang.

———, 2000b. « The Structural and Functional Composition of Romani Demonstratives ». *Grammatical Relations in Romani: The Noun Phrase*, édité par Viktor Elsik et Yaron Matras. 95-122.

———, 2002. *Romani: A Linguistic Introduction*. Cambridge : Cambridge University Press.

———, 2004. « Romacilikanes: The Romani Dialect of Parakalamos ». *Romani Studies* 14(1) : 59-109.

———, 2005a. « The Classification of Romani Dialects: A Geographical-Historical Perspective ». *General and Applied Romani Linguistics*, édité par Gerd Ambrosch, Barbara Schrammel et Dieter W. Halwachs. 7-32. Munich : Lincom.

———, 2005b. « The future of Romani: Toward a policy of linguistic pluralism ». *Roma Rights Quarterly* 1 : 31-44.

———, 2009. *Language Contact*. Cambridge : Cambridge University Press.

———, 2010. *Romani in Britain: the Afterlife of a Language*. Edinburgh : Edinburgh University Press.

———, 2011. « Romani ». *The Languages and Linguistics of Europe. A Comprehensive Guide*, édité par Bernd Kortmann et Johan van der Auwera. 257-68. Berlin : Mouton de Gruyter. http://romani.humanities.manchester.ac.uk/virtuallibrary/librarydb//web/files/pdfs/365/paper10.pdf

———, 2012. *A Grammar of Domari*. Berlin : Mouton De Gruyter.

———, 2013a. « Language and the Rise of a Transnational Romani Identity ». *RomIdent Working Papers.* http://romani.humanities.manchester.ac.uk/virtuallibrary/librarydb//web/files/pdfs/379/Paper24.pdf

———, 2013b. « Mapping the Romani Dialects of Romania ». *Romani Studies* 5. 23(2) : 199-243.

———, 2015. « Transnational policy and 'authenticity' discourses on Romani language and identity ». *Language in Society* 44 : 295-316.

Matras, Yaron et Peter Bakker, 1997. « Introduction ». *The Typology and Dialectology of Romani*, édité par Yaron Matras, Peter Bakker et Hristo Kyuchukov. vii-xxx. Amsterdam : John Benjamins Publishing Company.

Matras, Yaron et Viktor Elšík, 1998. *Romani Morpho-Syntax Database*. Romani Humanities Manchester. https://romani.humanities.manchester.ac.uk//rms/

Matras, Yaron et Jeanette Sakel, 2007. « Investigating the Mechanisms of Pattern Replication in Language Convergence ». *Studies in Language* 31 : 829-865.

Matras, Yaron, Christopher White et Viktor Elšík, 2009. « The Romani Morpho-Syntax (RMS) Database ». *The Use of Databases in Cross-Linguistic Studies*, édité par Martin Everaert, Simon Musgrave et Alexis Dimitriadis. 329-362. Berlin : Mouton de Gruyter.

Matushansky, Ora, 2005. « Les adjectifs - une introduction ». *Recherches Linguistiques de Vincennes* 34 : 9-54.

———, 2008. « On the Attributive Nature of Superlatives ». *Syntax* 11 : 26-90.

McNally, Louise et Christopher Kennedy, 2008. *Adjectives and Adverbs: Syntax, Semantics and Discourse.* Oxford : Oxford University Press.

Mecaj, Stela, 2016. « The Protection of the Roma Community's Rights and in Particular the Roma Children's Rights in Albania ». *Mediterranean Journal of Social Sciences* 7(3. S1) : 187-191.

Menantaud, Henri. 2011. « Forme adnumérale et indéfinitude en letton et en polonais ». *Baltic Languages, Slavic Languages: Contact, Confrontation, Comparison*, édité par Daniel Petit et Claire Le Feuvre. 281-288. Paris : CNRS Editions.

———, 2015. « Sur quelques emplois des désinences dites définies de l'adjectif letton ». Présentation à la journée d'études 'Modes et degrés d'individuation dans la construction des expressions référentielles', Inalco, Paris, 21 novembre.

Meyerhoff, Miriam, 2011. *Introducing Sociolinguistics*. 2ème édition. London & New York : Routledge.

Miklosich, Franz, 1872. *Über die Mundarten und Wanderungen der Zigeuner Europas*. Wien : Karl Gerold's Sohn.

Mladenova, Olga M., 2007. *Definiteness in Bulgarian: Modelling the Processes of Language Change*. Berlin : Mouton de Gruyter.

Moravcsik, Edith, 2003. « Non-Compositional Definiteness Marking in Hungarian Noun Phrases ». *Noun Phrase Structure in the Languages of Europe*, édité par Frans Plank. 397-466. Berlin & New York : Mouton de Gruyter.

Morzycki, Marcin, 2015. *Modification.* Cambridge : Cambridge University Press.

Mürvet, Enç, 1991. « The Semantics of Specificity ». *Linguistic Inquiry* 22(1) : 1-25.

Newman, W. L., 1969. « Gypsies in Albania ». *Journal of the Gypsy Lore Society*, Third Series 48(1–2) : 73.

Nølke, Henning, 2001. *Le regard du locuteur : pour une linguistique des traces énonciatives, vol.2.* Paris : Kimé.

Ntelitheos, Dimitrios, 2004. *Syntax of Elliptical and Discontinuous Nominals*. Thèse de master, Los Angeles : UCLA.

O'Rourke, Bernadette et Fernando F. Ramallo, 2011. « The Native-Non-Native Dichotomy in Minority Language Contexts: Comparisons between Irish and Galician ». *Language Problems & Language Planning* 35(2) : 139-159.

———, 2013. « Competing Ideologies of Linguistic Authority amongst "New Speakers" in Contemporary Galicia ». *Language in Society* 42(3) : 287-305.

Pădure, Cristian, 2019. *La variation copule/clitique sujet en romani du Mexique au contact de ser et estar en espagnol*. Thèse de doctorat, Paris : Institut National des Langues et Civilisations Orientales.

Pădure, Cristian, Stefano De Pascale, et Evangelia Adamou. 2018. « Variation between the copula *si* 'to be' and the *l*-clitics in Romani spoken in Mexico ». *Romani Studies* 28 (2): 263-92.

Panagiotidis, Phoevos, 2000. « Demonstrative Determiners and Operators: The Case of Greek ». *Lingua* 110 : 717-742.

———, 2005. *Determiner Spreading as DP-Predication*. Cyprus College.

Panagiotidis, Phoevos et Theodoros Marinis, 2011. « Determiner Spreading as DP-Predication ». *Studia Linguistica* 65(3) : 268-298.

Partee, Barbara, 1995. « Lexical Semantics and Compositionality ». *An Invitation to Cognitive Science*, édité par Lila Gleitman et Mark Liberman. Cambridge, Massachusetts : MIT Press.

Paspati, Alexandros Georgiou, 1870. *Études sur les Tchinghianés ou Bohémiens de l'Empire ottoman*. Constantinople : A. Koroméla.

Pat-El, Na'ama, 2009. « The Development of the Semitic Definite Article: A Syntactic Approach ». *Journal of Semitic Studies* LIV(1) : 19-50.

Penanros, Hélène de, 2015. « Les adjectifs complexes du lituanien ». Présentation au Séminaire du SeDyL 'Théories et données linguistiques', Inalco, Paris.

Piasere, Leonardo, 1982. « La terminologie des parents consanguins chez deux groupes de Roms ». *Études Tsiganes* 28(2) : 1-24.

———, 2011. *Roms : une histoire européenne*. Paris : Bayard.

Plank, Frans. 1995. « (Re-)Introducing Suffixaufnahme ». *Double Case: Agreement by Suffixaufnahme*, édité par Frans Plank. 3-110. Oxford & New York: Oxford University Press.

———, 2003a. *Noun Phrase Structure in the Languages of Europe*. Berlin : Mouton de Gruyter.

———, 2003b. « Double Articulation ». *Noun Phrase Structure in the Languages of Europe*, édité par Frans Plank. 337-395. Berlin : Mouton de Gruyter.

Plank, Frans et Edith Moravcsik, 1996. « The Maltese Article: Language-Particulars and Universals ». *Rivista di Lingüística* 8(1) : 183-212.

Poitou, Jacques, 2010. « Fiches de grammaire allemande ». Université Lumière Lyon II, cours de 2001-2010. http://j.poitou.free.fr/pro/html/div/fiches.html

Prévost, Sophie, 2003. « Les compléments spatiaux : du topique au focus en passant par les cadres ». *Travaux de linguistique* 2(47) : 51-77.

Queixalós, Francesc, 2000. *Syntaxe sikuani*. Paris : Peeters.

Rappaport, Gilbert C., 2014. « Determiner Phrases and Mixed Agreement in Slavic ». *Nominal Constructions in Slavic and Beyond*, édité par Lilia Schürcks, Anastasia Giannakidou, Urtzi Etxeberria et Peter Kosta. 343-390. Berlin & Boston : Mouton de Gruyter.

Reinhart, Tanya, 1982. *Pragmatics and Linguistics. An Analysis of Sentence Topics*. Bloomington : Indiana University Linguistics Club.

———, 2006. *Interface Strategies*. Cambridge, Massachusetts : MIT Press.

Riegel, Martin, Jean-Christophe Pellat et René Rioul, 2009. *Grammaire méthodique du français.* Paris : Presses universitaires de France.

Rivero, Maria-Luisa et Angela Ralli, éd., 2001. *Comparative Syntax of Balkan Languages*. Oxford : Oxford University Press.

Rizzi, Luigi, 1997. « The Fine Structure of the Left Periphery ». *Elements of Grammar*, édité par Liliane Haegeman. 281-339. Dordrecht : Kluwer.

Romani Project. *Grouping the Dialects of Romani. Dialect Classification.* http://romani.humanities.manchester.ac.uk/whatis/classification/dialect_classify.shtml

Rotaru, Julieta, 2016. « Peculiarities of the Romani Language Spoken in Wallachia and Moldavia in the 19th Century ». Présentation à la 12th International Conference on Romani Linguistics, Södertörn University, Stockholm, 16 septembre.

———, 2018. « Barbu Constantinescu, the first Romanian scholar of Romani studies ». *Romani Studies* 28(1) : 41-78.

Rubin, Aaron D., 2005a. *Studies in Semitic Grammaticalization*, vol.57. Harvard Semitic Studies. Winona Lake, Indiana : Eisenbrauns.

———, 2005b. « Definite Articles ». *Studies in Semitic Grammaticalization*. Harvard Semitic Studies. Winona Lake, Indiana : Eisenbrauns.

Runić, Marija, 2012. « The 'Definite Article' in the Slovene Dialect of Resia ». *Quaderni di lavoro ASIt* 16 : 91-106.

———, 2015. « The Definite Article in an Articleless Language ». *Slavic Grammar from a Formal Perspective. The 10th Anniversary FDSL Conference, Leipzig 2013*, édité par Gerhild Zybatow, Petr Biskup, Marcel Guhl et al. 411-425. Peter Lang.

Sakel, Jeanette et Daniel L. Everett, 2012. *Linguistic Fieldwork*. Cambridge : Cambridge University Press.

Samara, Miço, 1990. « Note de lecture : Stuart-E. Mann "A Vocabulary of Albanian Romani" ». *Études Tsiganes* 4 : 61.

Sampson, John, 1926. *The Dialect of the Gypsies Of Wales, Being the Older Form of British Romani Preserved in the Speech of the Clan of Abram Wood*. Oxford : The Clarendon Press.

Samvelian, Pollet, 2017. « Comment rendre compte de la variation en syntaxe ? Méthodes empiriques et quantitatives ». Séminaire doctoral Paris, mars. http://www.univ-paris3.fr/comment-rendre-compte-de-la-variation-en-syntaxe-methodes-empiriques-et-quantitatives-236120.kjsp

Sandfeld, Kristian, 1930. *Linguistique balkanique : problèmes et résultats*. Paris : Honoré Champion.

Sankoff, Gillian, 2005. « Cross-Sectional and Longitudinal Studies in Sociolinguistics ». *Sociolinguistics/Soziolinguistik: An International Handbook of the Science of Language and Society/Ein internationales Handbuch zur Wissenschaft von Sprache und Gesellschaft*, 2ème édition, édité par Ulrich Ammon, Norbert Dittmar, Klaus J. Mattheier et Peter Trudgill. 1003-1013. Berlin & New York : Walter de Gruyter.

———, 2006. « Age: Apparent Time and Real Time ». *Elsevier Encyclopedia of Language and Linguistics*, édité par Keith Brown. Oxford : Elsevier.

Sankoff, Gillian et Hélène Blondeau, 2007. « Language Change across the Lifespan: /R/ in Montreal French ». *Language* 83(3) : 560-588.

Sarău, Gheorghe et Ionel Cordovan, 2011. *Limba şi literatura rromani - manual pentru anul X de studiu*. Bucureşti: Editura Vanemonde.

Schmidt, Johannes, 1872. *Die Verwandtschaftsverhältnisse der indogermanischen Sprachen*. Weimar : H. Böhlau.

Schoorlemmer, Erik, 2012. « Definiteness Marking in Germanic: Morphological Variations on the Same Syntactic Theme ». *The Journal of Comparative Germanic Linguistics* 15(2) : 107-156.

Schroeder, Christophe, 2006. « Articles and Article Systems in Some Languages in Europe ». *Pragmatic Organization of Discourse in the Languages of Europe*, édité par Giuliano Bernini. Berlin : Mouton de Gruyter.

Schütze, Carson T., 1996. *The Empirical Base of Linguistics : Grammaticality Judgements and Linguistics Methodology*. Chicago : University of Chicago Press.

Scott, Gary-John, 1998. « Stacked Adjectival Modification and the Structure of Nominal Phrases ». *SOAS Working Papers in Linguistics and Phonetics* 8 : 59-89.

Sechidou, Eirene, 2011. *Balkan Romani: The Dialect of Ajios Athanasios-Greece*. München : LINCOM Europa.

Shanaj, Hera, Elda Hallkaj et Chris Cuninghame, 2015. *Analyse de situation des enfants en Albanie.* Pays-Bas : UNICEF Pays-Bas, Belgique, Suède.

Simon, Patrick, Emira Galanxhi et Olgeta Dhono, 2015. *Roma and Egyptians in Albania: A Socio-Demographic and Economic Profile Based on the 2011 Census*. United Nations Programme Support to Social Inclusion Initiative in Albania. Tiranë : Swiss Agency for Development and Cooperation. http://www.al.undp.org/content/dam/albania/docs/Census%202011%20Profile%20of%20Roma%20and%20Egyptians%20final.pdf

Skopeteas, Stavros, 2012. « The Empirical Investigation of Information Structure ». *The Expression of Information Structure*, édité par Manfred Krifka et Renate Musan. 216-246. Berlin : Mouton de Gruyter.

Skopeteas, Stavros et Caroline Féry, 2007. « Contrastive Topics in Pairing Answers: A Cross-Linguistic Production Study ». *Linguistic Evidence 2006*, édité par Sam Featherston et Wolfgang Sternefeld. 327-347. Berlin : Mouton de Gruyter.

Skopeteas, Stavros, Caroline Féry et Rusudan Asatiani, 2015. « Prosodic Phrasing in Scripted and Unscripted Speech - An Empirical Study on Georgian Prosodic Structure ». *Journal of Phonetics* : 1-40.

Skopeteas, Stavros, Ines Fiedler, Sam Hellmuth et al., 2006. « Questionnaire on Information Structure (QUIS): Reference Manual ». *Interdisciplinary Study on Information Structure*, 4. Potsdam : Audiovisuelles Zentrum der Universität Potsdam und Satz Druck GmbH Potsdam.

Sobolev, Andrej N., 2004. « On the Areal Distribution of Syntactic Properties in the Languages of the Balkans ». *Balkan Syntax and Semantics*, édité par Olga Mišeska Tomić. 59-100. Amsterdam & Philadelphia : John Benjamins Publishing Company.

Soravia, Giulio, 1972. « Italian Influences on the Dialect of the Gypsies of Abruzzi ». *Journal of the Gypsy Lore Society*, Third Series (51) : 34-39.

Spike, Gildea, 2014. « Working with Speakers: Comparative and Grammatical Elicitation, Experimental Stimuli, Orthographies and Texts ». Présentation au Séminaire du SeDyL 'Pratiques langagières - terrains, méthodes, théories', CNRS, Villejuif, 1er octobre.

Sproat, Richard et Chilin Shih, 1988. « Prenominal Adjectival Ordering in English and Mandarin ». *Proceedings of NELS 18*, édité par James Blevins et Juli Carter. 465-489. Amherst, Massachussetts : GLSA.

———, 1990. « The Cross-Linguistics Distribution of Adjectival Ordering Restrictions ». *Interdisciplinary Approaches to Language: Essays in Honor of S-Y. Kuroda*, édité par Carol P. Georgopoulos et Roberta Ishihara. 565-593. Dordrecht : Kluwer Academic Publishers.

Stavrou, Melita, 1991. « Nominal Apposition: More Evidence for the DP Analysis of NP ». *Empirical Approaches to Language Typology*, édité par John Payne. Berlin & New York : Mouton de Gruyter.

———, 1996. « Adjectives in Modern Greek: An Instance of Predication or an Old Issue Revisited ». *Journal of Linguistics* (32) : 79-112.

———, 1999. « The Position and Serialization of APs in the DP: Evidence from Greek ». *Studies in Greek Syntax*, édité par Artemis Alexiadou, Geoffrey Horrocks et Melita Stavrou. 201-225. Dordrecht : Kluwer Academic Publishers.

———, 2012. « Postnominal Adjectives in Greek Indefinite Noun Phrases ». *Functional Heads: The Cartography of Syntactic Structures*, édité par Laura Brugé, Anna Cardinaletti, Giuliana Giusti, Nicola Munaro et Cecilia Poletto. 379-393. Oxford : Oxford University Press.

Stivers, Tanya et Nicholas J. Enfield, 2010. « A Coding Scheme for Question-Response Sequences in Conversation ». *Journal of Pragmatics* 42(10) : 2620-2626.

Stojanovski, Aleksandar, 1989. « Ромите на балканскиот полуостров (врз основа на еден извор од 1523 г.) ». *Македонија во турското средновековие: од крајот на XIV почетокот на XVIII век*. 127-180. Skopje : Kultura.

Sula Raxhimi, Enkelejda, 2015. *Une ethnographie de la gestion de la population rom en Albanie*. Thèse de doctorat, Montréal : Université de Montréal. https://papyrus.bib.umontreal.ca/xmlui/handle/1866/13958

Szendrői, Kriszta, 2010. « A Flexible Approach to Discourse-Related Word Order Variations in the DP ». *Lingua* 120 : 864-878.

Tagliamonte, Sali A. et R. Harald Baayen, 2012. « Models, Forests, and Trees of York English: *Was/Were* Variation as a Case Study for Statistical Practice ». *Language Variation and Change* 24(02) : 135-178.

Taube, Moshe 2012. « On the Relative Marker Vos and Yiddish Postmodifiers ». *Leket: Yidishe Shṭudyes Haynṭ = Jiddistik Heute = Yiddish Studies Today* 1 : 467-481.

Tenser, Anton, 2005. *Lithuanian Romani*. München : Lincom Europa.

Thomas, Paul-Louis et Vladimir Osipov, 2012. *Grammaire du bosniaque croate monténégrin serbe*. Paris : Institut d'Études Slaves.

Thuilier, Juliette, 2012. *Contraintes préférentielles et ordre des mots en français.* Université Paris VII Diderot.

———, 2015. « La position de l'adjectif en français : approche quantitative sur corpus et expériences psycholinguistiques ». Présentation au Séminaire ERSS, Université Toulouse - Jean Jaurès, Toulouse, 26 mars.

Tirard, Aurore, 2014. « Nominal Overdetermination in Romani ». Présentation à la 11th International Conference on Romani Linguistics, University of Oslo, Oslo, 16 septembre.

———, 2015. « A Variationist Approach to Syntactic Doubling: the Case of Romani ». Présentation à l'International Conference on Language Variation in Europe 8, Universität Leipzig, Leipzig, 27 mai. http://conference.uni-leipzig.de/iclave8/programme.pdf

———, 2017. « Syntactic Doubling and Variation: The Case of Romani ». *Language Variation: European Perspectives VI. Selected Papers from the Eighth International Conference on Language Variation in Europe (ICLaVE 8)*, édité par Isabelle Buchstaller et Beat Siebenhaar. 141-156. Amsterdam & Philadelphia: John Benjamins Publishing Company.

Tomić, Olga Mišeska, 2004. « The Balkan Sprachbund Properties ». *Balkan Syntax and Semantics*, édité par Olga Mišeska Tomić. 1-55. Amsterdam & Philadelphia : John Benjamins Publishing Company.

———, 2006. *Balkan Sprachbund Morpho-Syntactic Features*. Dordrecht : Springer.

Trovesi, Andrea, 2004. *La genesi di articoli determinativi - modalità di espressione della definitezza in ceco, serbo-lusaziano e sloveno*. Milano : FrancoAngeli.

Trubeckoj, Nikolaj Sergeevič, 1928. « Proposition 16 ». *Actes du premier congrès international de linguistes, La Haye, 10-15. Avril 1928*. Leiden : A. W. Sijthoff.

Uhlik, Rade, 1941. « Bosnian Romani: Vocabulary ». *Journal of the Gypsy Lore Society*, Third Series 20(3–4) : 100-140.

———, 1942a. « Bosnian Romani: Vocabulary (G-L) ». *Journal of the Gypsy Lore Society*, Third Series 21(1–2) : 24-54.

———, 1942b. « Bosnian Romani: Vocabulary (M-R) ». *Journal of the Gypsy Lore Society*, Third Series 21(3–4) : 110-140.

———, 1943a. « Bosnian Romani: Vocabulary (S-S) ». *Journal of the Gypsy Lore Society*, Third Series 22(1–2) : 38-46.

———, 1943b. « Bosnian Romani: Vocabulary (T-Z) ». *Journal of the Gypsy Lore Society*, Third Series 22(3–4) : 107-118.

———, 1944. « Five Tales from Bosnia ». *Journal of the Gypsy Lore Society*, Third Series 23 : 97-106.

———, 1956. *Iz ciganske onomastike [Über die Namen einiger Zigeunerstämme und deren Mundarten]*. Sarajevo : Veselin Masleša.

UNDP, Albania, 2012. *Roma and Egyptians in Albania*. Tirana : UNDP/World Bank/EC. http://www.al.undp.org/content/dam/albania/docs/misc/Roma%20in%20%20Albania%20June%202012%20profile.pdf

Vallduví, Enric, 1992. *The Informational Component*. New York : Garland Publishers.

Vaux de Foletier, François de, 1981. « Notes de lecture dans les revues ». *Études Tsiganes* 3 : 43.

Vencel', Tatjana V. et Lev N. Čerenkov, 1976. « Dialekty cyganskogo jazyka ». *Jazyki Azii i Afriki*. 283-332. Moskva : Nauka.

Verhoeven, Elisabeth et Stavros Skopeteas, 2015. « Licensing Focus Constructions In Yucatec Maya ». *International Journal of American Linguistics* 81(1) : 1-40.

Wallisch, Friedrich, 1931. *Neuland Albanien.* Stuttgart : Franckh'sche Verlagshandlung.

Waltereit, Richard et Ulrich Detges, 2008. « Syntactic Change from within and from without Syntax: A Usage-Based Analysis ». *The Paradox of Grammatical Change – Perspectives from Romance*, édité par Ulrich Detges et Richard Waltereit. 13-30. Amsterdam & Philadelphia : John Benjamins Publishing Company.

Weigand, Gustav, 1924. *Ethnographie von Makedonien*. Leipzig : F. Brandstetter.

Wilmet, Marc, 1981a. « Antéposition et postposition de l'épithète qualificative en français contemporain ». *Travaux de Linguistique* 7 : 179-201.

———, 1981b. « La place de l'épithète qualificative en français contemporain. Étude grammaticale et stylistique ». *Revue de Linguistique Romane* 45 : 17-73.

Winford, Donald, 2003. *An Introduction to Contact Linguistics*. Oxford : Blackwell Publishing.

Winter, Yoad, 2004. « On Some Problems of (in)Definiteness within Flexible Semantics ». *Lingua* 115 : 767-786.

Yatziv-Malibert, Il-Il. 2015. « Défini et non défini en hébreu contemporain ». Présentation au Séminaire du SeDyL 'Théories et données linguistiques', Inalco, Paris, 9 janvier.

Zatreanu, Mihaela et Dieter W. Halwachs, 2003. *Romani in Europe*. Strasbourg : Council of Europe.

Zavitziano, Spiridion C., 1901. « TURKEY. Report from Constantinople — Plague in That City ». *Public Health Reports (1896-1970)* 16(42) : 2458-2462.

Zavitziano, Spiridion C., Dr. Tachdjion, Dr. Xanthropaulides, Dr. Pompouras et Dr. Cozzonis 1900. « TURKEY. Plague in Djivanro ». *Public Health Reports (1896-1970)* 15(18) : 1098-1101.

Zlatić, Larisa, 2014. « Definiteness and Structure of NPs in Slavic ». *Nominal Constructions in Slavic and Beyond*, édité par Lilia Schürcks, Anastasia Giannakidou, Urtzi Etxeberria et Peter Kosta. 17-38. Berlin & Boston : Mouton de Gruyter.

Annexes

1. Annexes du chapitre I : Les Roms d'Albanie et leur langue

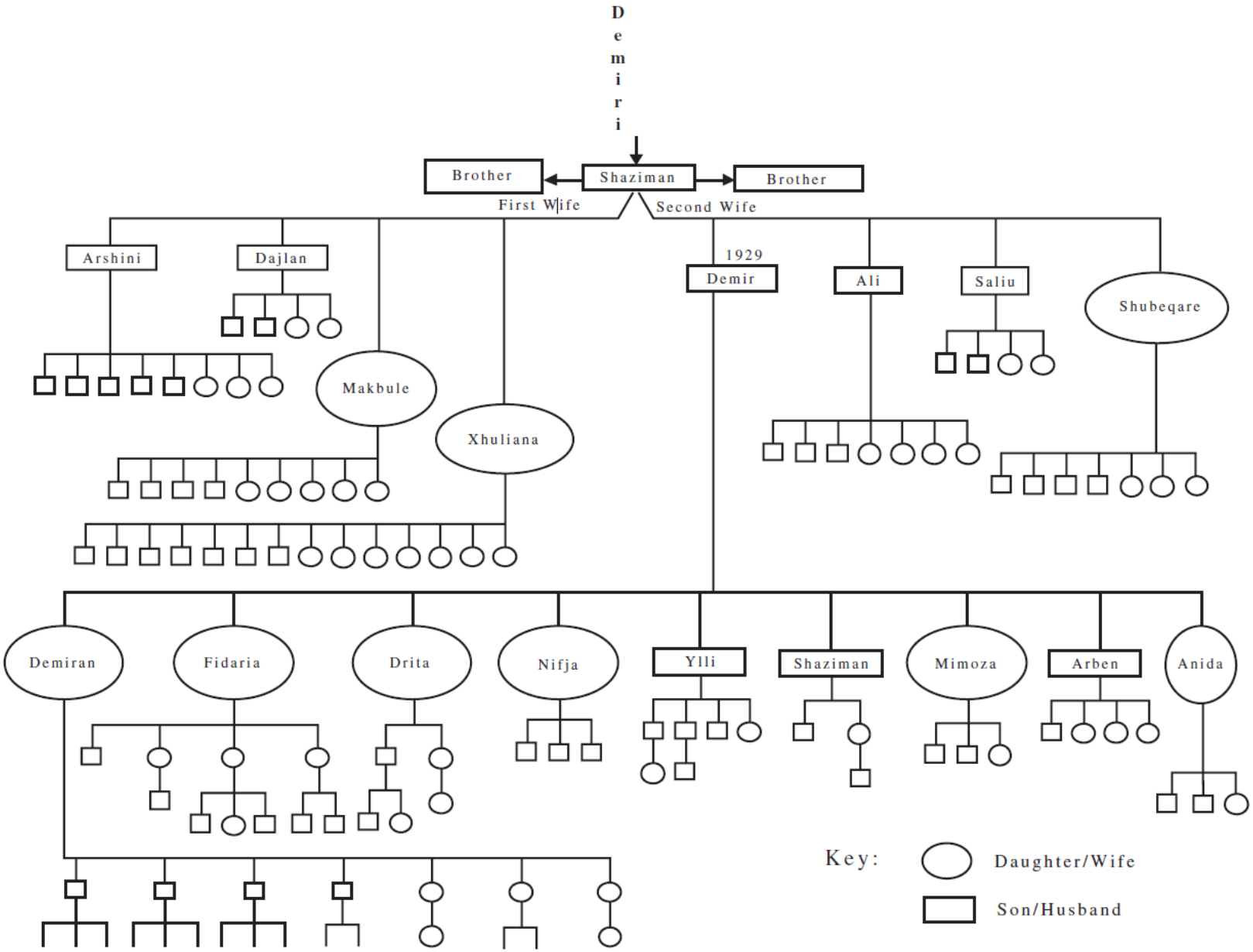


Figure 88 : ***The Patrilineal Kinship Structure of One Roma Fis*** **(De Soto, Beddies et Gedeshi 2005 : 222)**

Structure of Arben's *Fis* in Korça

"My grandfather's name is Shaziman. He was born about 100 years ago in Kostur, Greece, as the son of Demir Demirit, and he had two wives. My grandfather had four children with his first wife—two boys [Arshin and Dajlan] and two girls [Makbule and Xhuliana]. He had four children with his second wife—three boys [Demir, Ali, Sali] and one girl [Shubeqare]. The oldest son from his first wife, Arshin, had nine children, while the other son, Dajlan, had four children. The oldest daughter, Makbule, was married in Korça and had eight children. The other daughter from his first wife, Xhuliana, was married in a village in Fier [Zhupan] and had 14 children.

My father, Demir, had his grandfather's name. He was the first son from my grandfather's second wife and had nine children—three boys and six girls. The second son, Ali, had three boys and our girls. So, just taking the first cousins from my father's side, we have over 60 people. Each one of them is married and has three or four children. Just from my father Demir's side, there are 94 of us, whereas Makbule with her grandchildren numbered 120. So, starting with my grandfather, Shaziman, there are a few hundred of us. The *fis* is made up of all cousins that have my last name, Demir." Quoted from a male Roma respondend.

Figure 89 : *Structure of Arben's Fis in Korça* (De Soto, Beddies et Gedeshi 2005 : 22)

Function of Arben's *Fis* in Korça

"They say that my grandfather, Shaziman, was very well-off and had a lot of influence with other Roma. In King Zog's time, if someone was arrested by the police without cause, he would intervene and have the man freed. He resolved many conflicts between Roma families and because of this, he was referred to by all as Shaziman Aga. He was the only Roma from Korça to have gone to America three times as a migrant. He worked there in metallurgy, but in addition to this he was involved in horse dealing. My grandfather had two brothers from different mothers and one of them lived in the village of Çuka in Saranda. The other migrated to America and later returned to Greece, where he died. Shaziman was half-Karbuxh Roma on his father's side, and half-Cergar Roma on his mother's side, but he lived among the Karbuxh fis of Korça.

My father, Demir, was a very able person and had a lot of influence within the Roma community. He maintained relations with the local government authorities, arranged marriages, and resolved conflicts. He inherited this influence partly from his father and partly from his own achievement. He was involved in private business, and he continued to do business during the socialist period, when private enterprise was strictly forbidden." Quoted from a male Roma respondend.

Figure 90 : *Function of Arben's Fis in Korça* (De Soto, Beddies et Gedeshi 2005 : 23)

2. Annexes du chapitre IV : Analyse du corpus spontané

2.1. Profil des locuteurs ayant prononcé les exemples spontanés

Tableau 114 : Profil des locuteurs ayant prononcé les exemples spontanés

Nom	Age[1]	Genre	Lieu d'enreg.	Lieu de naissance / de résidence[2]	Variété	Autres langues connues[3]	Nombre années scolarité	Profession	Militantisme
F16AČ	16	F	Korçë	Korçë	arli & čergar I	albanais	6 à 12 années	écolière	non
M18Č	18	M	Korçë	Fier	čergar I	albanais (+ anglais)	6 à 12 années	écolier	non
F22M	22	F	Llakatund	Tiranë	mečkar	albanais (+ anglais, italien)	plus 12 années	étudiante	oui
M22A	22	M	Korçë	Korçë	arli	albanais	0 à 5 années	artisan menuisier	non
M23 M	23	M	Llakatund	Llakatund	mečkar	albanais (+ italien)	6 à 12 années	étudiant	non
F24A	24	F	Bilisht	Pogradec / Bilisht	arli	albanais, (+ ?)	0 à 5 années (0)	sans emploi	non
M24M	24	M	Llakatund	Berat	mečkar	albanais (+ grec)	6 à 12 années	enseignant de romani	oui
M27A	27	M	Korçë	Korçë	arli	albanais, turc	0 à 5 années (0)	artisan	non
F38A	38	F	Korçë	Tiranë / Korçë	arli	albanais	0 à 5 années (0)	sans emploi	non
F44AČ	44	F	Korçë	Fier / Korçë	čergar I & arli	albanais, grec	0 à 5 années (0)	femme au foyer	non
M46A	45 (2013) 46 (2014)	M	Korçë	Korçë	arli	albanais (+ grec)	6 à 12 années	journaliste	oui
F47M	47	F	Llakatund	Elbasan / Llakatund	mečkar	albanais	0 à 5 années	femme au foyer	non
M50A	50	M	Korçë	Korçë / Elbasan, Athènes, Korçë	arli	albanais, grec	6 à 12 années	sans emploi	non
M58MČ	58 (2013)	M	Voskopojë	Tiranë	mečkar & čergar I	albanais, (+ serbe)	plus 12 années	chorégraphe, enseignant de langue et culture romani, à la retraite	oui
F60A	60	F	Korçë	Korçë	arli	albanais (+ grec, + turc)	0 à 5 années	vendeuse à la retraite	non
F65A	65	F	Korçë	Korçë / Florina	arli	albanais, grec	0 à 5 années	agricultrice puis femme au foyer et retraitée	non
F66M	66	F	Voskopojë	Fier / Vlorë, Korçë, Berat, Tiranë	mečkar	albanais	6 à 12 années	enseignante de romani et militante	oui
M68A	68 (2013)	M	Voskopojë	Tiranë	arli	albanais, (+ ?)	6 à 12 années	éducateur pour enfants, militant, puis à la retraite	oui
M71A	70 (2013) 71 (2014)	M	Korçë	Rembesi / Florina	arli	albanais, grec (+ macédonien)	0 à 5 années	maçon puis vendeur, à la retraite	non
F71A	71 (2014)	F	Bilisht	Bitinckë / Bilisht	arli	albanais	0 à 5 années	femme au foyer puis retraitée	non

[1] A l'époque de l'enregistrement (2013 ou 2014).

[2] Nous mentionnons derrière la barre oblique le lieu de résidence si celui-ci est ou a été différent du lieu de naissance.

[3] La mention des langues connues autres que le romani correspond aux pratiques déclarées des locuteurs. J'ai distingué deux niveaux de connaissance : plutôt fluide ou bien plutôt partielle (indiquée entre parenthèses).

2.2. De la polysémie de certains termes de famille

L'exemple (167) appelle des précisions. *Čhavo* et *Rom* sont des termes polysémiques, qui entrent en opposition avec des termes différents en fonction du sème focalisé dans tel ou tel emploi en contexte. Le détail est présenté en (Tableau 115).

Tableau 115 : Analyse des sèmes de *čhavo* et de *rom*

čhavo			***rom***		
Sème	**Traduction**	**Antonyme**	**Sème**	**Traduction**	**Antonyme**
Ethnie	garçon rom	*raklo*, garçon non-rom	Ethnie	homme rom	*gadžo*, homme non-rom
Age	enfant	*rom*, adulte	Age	adulte	*čhavo*, enfant
Genre	garçon	*čhaj*, fille	Genre	homme	*romni*, femme
Relation familiale	fils	*čhaj*, fille & *o dad i daj*, parents	Relation familiale	époux	*romni*, épouse

Les diverses interprétations de l'exemple (167) illustrent la polysémie de *čhavo*. La suite du conte dont est issu l'exemple (158) illustre quant à lui la polysémie de *Rom*. Le personnage principal, *o phral o baro*, le grand frère, dont on sait qu'il n'est pas Rom (*o baro raklo*, le grand garçon non-rom), a abandonné sa femme enceinte. Lorsqu'il revient vingt ans plus tard, il se retrouve père d'un jeune homme par définition non-rom lui aussi. Le locuteur (arli) dit : « *Akana o baro phral/ kan džala ćhere/ o biše brešengo čhavo ondilo/ Rom*. », « Maintenant le frère aîné/ quand il rentre chez lui/ son fils de vingt ans est devenu/ un homme. » On voit qu'il n'y a pas de contradiction à ce qu'un *raklo*, enfant non-rom, devienne un jour *Rom*, au sens d'adulte et non de Rom. Peut-être est-ce pour éviter cette ambiguïté que la variété arli a emprunté à une langue encore indéterminée le terme *maksemi* pour désigner l'enfant, avec les seuls sèmes de l'âge (jeune enfant opposé à adulte) et de la relation familiale (enfant opposé à parent).

2.3. Analyses

Synthèse des analyses présentées pour les exemples polydéfinis et monodéfinis, pour les exemples polydéfinis et monodéfinis avec rupture interne :

Tableau 116 : Synthèse des analyses présentées pour les exemples polydéfinis

Structure	n°	Syntagme	Statut de l'information	Structure de l'information	Interprétation sémantique
DEF N DEF A	**(154)**	*o phral o vaver*	accessible (A nouveau)	focus contrastif	A – restriction, contraste *in presentia*, propriété spécifique
	(155)	*mo phral o xurdo*	accessible (A nouveau)	topique-propos	A – restriction, contraste *in presentia*, propriété spécifique
	(156)	*mi phen i/ i tikni*	donné (A nouveau)	topique-propos	A – restriction, contraste *in presentia*, propriété spécifique
	(157)	*mo phral o tikno*	donné (A donné)	topique-propos	A – contraste *in presentia*, propriété spécifique
	(158)	*o phral o baro*	donné (A donné)	topique-propos	A – restriction, contraste *in presentia*, info présentée comme nouvelle
	(159)	*i čhaj i vaver*	donné (A nouveau)	topique-propos	A – restriction, contraste *in presentia*, propriété spécifique
	(160)	*i škola i bari*	donné (A nouveau)	topique-cadre	A – ? restriction, contraste, ou propriété inhérente et spécifique
	(161)	*o phral o tikno*	donné (A donné)	topique-propos	A – restriction, contraste *in presentia*, propriété spécifique emphatisée
	(162)	*o filme o bare*	nouveau (A nouveau)	focus nouvelle info non sollicitée	A – restriction, contraste *in absentia*, propriété spécifique emphatisée..
	(163)	*i bira i purani*	accessible (A nouveau)	focus nouvelle info non sollicitée	A – contraste *in absentia*, propriété spécifique emphatisée
	(164)	*o lokali o baro*	accessible (A nouveau)	focus nouvelle info non sollicitée	A – restriction, contraste *in presentia*, propriété spécifique emphatisée
	(165)	*o lokal o baro*	accessible (A nouveau)	topique-cadre	A – restriction, contraste *in absentia*, propriété spécifique emphatisée.
	(166)	*amaro ćher o baro*	accessible (A donné)	focus nouvelle info sollicitée	A – restriction, contraste *in presentia*, propriété spécifique emphatisée
	(167)	*o čhave o bare*	accessible (A nouveau)	focus contrastif	A – restriction, contraste *in presentia*, propriété spécifique emphatisée
	(168)	*to ćidipe to/ gadžikane*	accessible (A nouveau)	focus nouvelle info non sollicitée	A – restriction, contraste *in absentia*, propriété spécifique emphatisée
	(169)	*o problemi o baro*	accessible (A nouveau)	topique-propos	A – restriction, contraste *in presentia*, propriété spécifique emphatisée
DEF A DEF N	**(261)**	*o tikne o maksemi*	donné (A nouveau)	topique-propos	A – ? restriction, contraste, ou propriété inhérente, référence globale
	(171)	*o barvalo o Skender*	accessible (A nouveau)	topique-propos	A – caractérisation, propriété inhérente et spécifique

Tableau 117 : Synthèse des analyses présentées pour les exemples monodéfinis

Structure	n°	Syntagme	Statut de l'information	Structure de l'information	Interprétation sémantique
DEF A N	**(172)**	*sa i romani čhib*	donné (A nouveau / accessible)	topique-propos	A – restriction, propriété non-spécifique
	(173)	*to purane breša*	nouveau (A nouveau)	topique-cadre	A – propriété non-spécifique
	(174)	*i tikni čhaj*	accessible (A nouveau)	topique-propos	A – restriction, propriété spécifique
	(175)	*kato baro Del*	nouveau (A nouveau / accessible)	topique-cadre	A – propriété inhérente
	(176)	*o šutalo thud*	donné (A donné)	focus contrastif	A – propriété inhérente, référence globale, générique
	(177)	*o bilačho manuš*	donné (A donné)	topique-propos	A – restriction, contraste implicite, propriété non-spécifique
	(178)	*o lače manuša*	nouveau (A nouveau)	topique-propos	A – restriction, contraste implicite, propriété non-spécifique
	(179)	*e bare doktorensar*	donné (A nouveau)	focus nouvelle info non sollicitée	A – propriété inhérente
	(188)	*o jekto kongreso*	nouveau (A nouveau)	topique-propos	A – propriété spécifique inhérente (contraste implicite)
	(189)	*o baro phral*	donné (A donné)	topique-propos	A – restriction, contraste implicite, propriété spécifique
	(190)	*o vaver phral*	donné (A nouveau)	topique-propos	A – restriction, contraste implicite, propriété spécifique
	(191)	*o maj purane Roma*	nouveau (A nouveau)[2]	topique-propos	A – restriction, contraste implicite, propriété spécifique
	(192)	*o vaver Roma*	accessible (A nouveau)	topique-propos	A – restriction, contraste implicite, propriété spécifique
DEF N A	**(180)**	*o pete bare*	nouveau (A nouveau)	focus nouvelle info non sollicitée	A – restriction, contraste implicite *in absentia*, propriété spécifique
	(181)	*e čhave bare*	accessible (A donné)	topique-propos	A – restriction, contraste *in presentia*, propriété spécifique
	(182)	*o čhavo cikno*	accessible (A donné)	topique-propos	A – restriction, contraste *in presentia*, propriété spécifique
	(183)	*to amalinja vavera*	nouveau (A nouveau)	focus nouvelle info non sollicitée	A – restriction, contraste *in absentia*, propriété spécifique,
	(184)	*odole gilja koraxane*	donné (A donné)	focus nouvelle info sollicitée	A – restriction, contraste *in presentia/absentia*, propriété spécifique
	(185)	*ano foro baro*	nouveau (A nouveau / accessible)	topique-cadre	A – restriction, contraste implicite *in presentia*, propriété spécifique

Structure	n°	Syntagme	Statut de l'information	Structure de l'information	Interprétation sémantique
	(186)	*pire gada romane*	nouveau (A nouveau)	focus nouvelle info non sollicitée	A –? restriction, contraste implicite *in absentia*, propriété non-spécifique
	(187)	*to vaćeribe řomano*	nouveau (A accessible)	focus contrastif	A – restriction, contraste *in presentia*, propriété spécifique

Tableau 118 : Synthèse des analyses présentées pour les exemples polydéfinis avec rupture interne

Structure	n°	Syntagme	Statut de l'information	Structure de l'information	Interprétation sémantique
DEF N/ DEF A	**(193)**	*to phral/ o barvalo*	donné (A donné) ou nouveau (A nouveau)	topique-propos	A – ? restriction, contraste *in absentia*, propriété spécifique emphatisée ou caractérisation, propriété inhérente et spécifique
	(194)	*me phrales/ e xurdes/ e Benes*	donné (A nouveau)	focus nouvelle info non sollicitée	A – restriction, contraste *in presentia*, propriété spécifique
	(195)	*o duj phrala Aurora/ o bare*	accessible (A nouveau)	topique-propos	A – ? restriction, contraste *in presentia*, propriété spécifique
	(196)	*e čhej kaj si ma i bari/*	donné (A nouveau)	topique-propos	A – ? restriction, contraste *in presentia*, propriété spécifique
	(197)	*i čhaj/ bučhola Nəɻtila i baɻi.*	donné (A nouveau)	topique-propos	A – restriction, contraste *in presentia*, propriété spécifique

Tableau 119 : Synthèse des analyses présentées pour les syntagmes à adjectif substantivé

Structure	n°	Syntagme	Statut de l'information	Structure de l'information	Interprétation sémantique
DEF N_1/ DEF N_2	**(198)**	*e makseme/ e/ e cekne*	donné (A nouveau)	focus nouvelle info non-sollicitée	A – propriété générique
DEM N	**(199)**	*odəja bari*	donné (A donné)	topique-propos	A – restriction, contraste, propriété spécifique
	(200)	*odja ave*	donné (A nouveau)	topique-propos	A – restriction, contraste, propriété spécifique

Tableau 120 : Synthèse des analyses présentées pour les exemples monodéfinis avec rupture interne

Structure	n°	Syntagme	Statut de l'information	Structure de l'information	Interprétation sémantique
DEF A/ N	**(201)**	*o baro/ doktori*	donné (A donné)	topique-propos	A – propriété inhérente et spécifique
DEF N/ A	**(202)**	*o pištolja/ purane*	nouveau (A nouveau)	focus nouvelle info non sollicitée	A – restriction, contraste *in absentia*, propriété spécifique
	(203)	*i tradita/ ə ˌomani*	accessible (A nouveau)	focus nouvelle info non sollicitée	A – restriction, contraste *in absentia*, propriété spécifique
	(204)	*i čhib ə/ gadžikani grećist*	nouveau (A nouveau + donné)	focus nouvelle info non sollicitée	A – restriction, contraste *in presentia*, propriété spécifique

2.4. Corpus spontané : DEF N DEF A

(154) F60A, Korçë, 21 juillet 2014

Cet échange est issu d'un entretien dirigé, où la locutrice présente sa famille et énumère ses frères et sœurs.

1	*F60A: Me daja čardena la Pakize.*
2	*M46A: Kećі maksemja isi te daja?*
3	*F60A: Me daja isi la enja maksemja.*
4	*M46A: Ka vaćeres mang lenđe lafja/ sare maksemenđe.*
5	*F60A: Mi/ mi bari phen isi Demirana.*
6	*M46A: Hmhm.*
7	*F60A: O/ odja vaver čardena la Fidarija.*
8	*E/ odəja i trin kaj si mi phen čardena la Nife.*
9	*Me isinom i štar/*
10	*M46A: štarenđiri.*
11	*F60A: E štarengi isinom me Drita čardena man.*
12	*Sora isi man daa jek vaver phen/ čardena la: Moza.*
13	*M46A: Hmhm.*
14	*F60A: Daa jek i tikni xurdi/ isi: Anida/ čardena la.*
15	*Sora isi mo baro phral/ kaj čardena les Yli.*
16	*M46A: Hmhm.*
17	*Isi o phral o vaver/ pala leste/ čardena les Šaziman thodam amare papusko lafi lesće.*
18	*M46A: šukar.*
19	*F60A: Ah/ isi man daa jek vaver phral/ čardena les Arben.*
20	*M46A: Te oves sasto baxtalo.*
1	F60A : Ma mère s'appelle Pakize.
2	M46A : Combien d'enfants a ta mère ?
3	F60A : Ma mère elle a neuf enfants.
4	M46A : Tu vas me dire leur nom/ à tous les enfants.
5	F60A : Ma/ ma grande soeur est Demirana.
6	M46A : Hmhm.
7	F60A : le/ l'autre s'appelle Fidaria de son prénom.
8	Euh/ la troisième qui est ma soeur elle s'appelle Nifa.
9	Moi je suis la quatre/
10	M46A : La quatrième.
11	F60A : La quatrième c'est moi, je m'appelle Drita.
12	Ensuite j'ai encore une autre sœur/ elle s'appelle : Moza.
13	M46A : Hmhm.
14	F60A : Encore une la cadette la benjamine/ c'est : Anida elle s'appelle.
15	Ensuite il y a mon grand frère/ qui s'appelle Yli.
16	M46A : Hmhm.
17	F60A : Il y a l'autre frère/ après lui/ il s'appelle Shaziman/ on lui a donné le nom de notre grand-père.
18	M46A : Bien.
19	F60A : J'ai encore un autre frère/ il s'appelle Arben.
20	M46A : Merci à toi.

(154)	*Isi*	***o***	***phral***	***o***	***vaver,***	*pala*	*les-te,*
	être.3.PRES	ART.DEF.M.SG	frère(M.SG)	ART.DEF.M.SG	autre(M.SG)	après	3.SG.OBL-LOC
	čard-ena	*les*	*Šaziman.*				
	appeler-3.PL.PRES	3.SG.OBL	Shaziman				

« Il y a un autre frère après lui, il s'appelle Shaziman. »

Statut de l'information

Le référent « l'autre frère » est accessible-inférable (*set relation*) puisque la locutrice a annoncé qu'elle avait huit frères et sœurs. L'adjectif *autre* est nouveau lexicalement et le nom *frère* donné-actif lexicalement.

Structure de l'information

Le syntagme *o phral o vaver*, l'autre frère, constitue le focus de la phrase, plus précisément un focus contrastif du type *subtype selection* chez Götze et al. (2007 : 180-181).

Interprétation sémantique

La partie nominale du syntagme, le frère, fait partie du contexte, du présupposé. La locutrice a en effet annoncé qu'elle avait de nombreux frères et sœurs et vient, dans la phrase qui précède, de mentionner le plus grand de ses frères (qui est en fait plus jeune qu'elle). Il s'agit ensuite de sélectionner un autre frère. Le contraste se fait donc *in presentia* en discours. Cette sélection se fait au moyen de l'adjectif, qui apporte un sème spécifique permettant d'identifier le référent. La référence se calcule par l'adjonction des sèmes de l'adjectif et de ceux du nom. En d'autres termes, le référent de l'expression globale est à l'intersection de l'ensemble {frères} et de l'ensemble {autres entités}. L'identification du référent se fait en opérant une restriction pragmatique, au sein des référents du nom grâce à l'adjectif.

(155) M18Č, Korçë, 24 juillet 2014

Cet échange est issu d'un entretien dirigé où le locuteur se présente.

1	*Ch: əəə/ save čhibja džanesa tu?*
2	*M18Č: Hmm?*
3	*Ch: Kećі/ kećі čhibja džanesa?*
4	*M18Č: Na/ me džanava ščip em romane.*
5	*Ch: Savo ə romane/ savi čhib romani džanesa?*
6	*M18Č: Čergarče.*
7	*Ch: Tiro dad em tiri daj isi čergarja?*
8	*M18Č: Va.*

9	*Ch: Em to papu em ti mami?*
10	*M18Č: Em mo papu em mi mami/ on si Kosova akana.*
11	*Bešena, kerena buti othar.*
12	*Ch: Kaj?*
13	*M18Č: Kosova.*
14	*Ch: Kaj? Kosova isi bari.*
15	*M18Č: Kosova!*
16	*Ch: Priština?*
17	*M18Č: Priština/ Priština.*
18	*Ch: I on čergarče džanena.*
19	*M18Č: džanena.*
20	*Ch: Aver čhibja džanena?*
21	*M18Č: Vaver čhib/ ščip em romane/ amari čhib/ čergarče.*
22	*Ch: ə/ grećišt džanesa?*
23	*M18Č: Na. ə/ mi daj džanela, kurse mo dat e džanela itališt.*
24	*Ch: Oh! But šukar. Em ə khorane khorane džanela?*
25	*M18Č: Va, džanel.*
26	*Ch: Em tu džanesa?*
27	*M18Č: Hmm/ mo phral o xurdo džanel anglišt.*
28	*Ch: Kat/ kati škola?*
29	*M18Č: Emm/ si klas e đasht. Em/ klas sa per? klas e đasht.*
1	Ch : Euh/ quelles langues tu parles toi ?
2	M18Č : Hmm ?
3	Ch : Combien/ combien de langues tu parles ?
4	M18Č : Non/ moi je parle albanais et romani.
5	Ch : Quel romani/ quelle langue romani tu parles ?
6	M18Č : Cergar.
7	Ch : Ton père et ta mère sont des Čergars ?
8	M18Č : Oui.
9	Ch : Et ton grand-père et ta grand-mère ?
10	M18Č : Et mon grand-père et ma grand-mère/ eux ils sont au Kosovo maintenant.
11	Ils habitent et travaillent là-bas.
12	Ch : Où ça ?
13	M18Č : Au Kosovo.
14	Ch : Où ? c'est grand/ le Kosovo.
15	M18Č : Au Kosovo !
16	Ch : A Priština ?
17	M18Č : A Priština/ à Priština.
18	Ch : Alors eux aussi parlent čergar.
19	M18Č : Oui.
20	Ch : Est-ce qu'ils savent d'autres langues ?
21	M18Č : Une autre langue/ l'albanais et le romani/ notre langue/ le čergar.
22	Ch : Euh/ ils parlent grec ?
23	M18Č : Non. Euh/ ma mère si, alors que mon père il parle l'italien.
24	Ch : Oh ! Super. Euh euh il parle le turc ?
25	M18Č : Oui/ il le parle.
26	Ch : Et toi aussi ?
27	M18Č : Hmm/ mon tout petit frère parle anglais.
28	Ch : Grâce/ grâce à l'école ?
29	M18Č : Hmm/ il est en 6e. Et/ en classe combien ? en sixième.

(155)	***M-o***	***phral***	***o***	***xurd-o***	*džan-el*	*angliš*t.
	POSS.1SG-M.SG	frère(M.SG)	ART.DEF.M.SG	tout_petit-M.SG	savoir-3SG.PRES	en_anglais
	« Mon petit frère parle anglais. »					

Statut de l'information

Le locuteur ayant déjà mentionné, en début d'entretien, qu'il a trois frères et sœurs, on peut considérer que le référent « mon tout petit frère » est accessible-inférable (*set relation*). L'adjectif *tout petit* est nouveau lexicalement et le nom *frère* est accessible-inférable dans la conversation, où d'autres termes de membres de la famille viennent d'être énoncés (grand-père, grand-mère, mère, père).

Structure de l'information

Le syntagme *mo phral o xurdo*, mon tout petit frère, constitue le topique-propos de l'expression *džanel angliš*t, parle anglais.

Interprétation sémantique

La partie nominale du syntagme, le frère, fait partie du contexte, du présupposé, puisque le locuteur a déjà mentionné qu'il avait trois frères et sœurs. Le contraste se fait donc *in presentia* en discours. Il s'agit de sélectionner un frère, sélection qui se fait au moyen de l'adjectif, qui apporte un sème spécifique permettant d'identifier le référent. La référence se calcule par l'adjonction des sèmes de l'adjectif et de ceux du nom. En d'autres termes, le référent de l'expression globale est à l'intersection de l'ensemble {frères} et de l'ensemble {toutes petites entités}. Le locuteur a un petit frère, une petite sœur, et un deuxième petit frère. Ce « deuxième » petit frère (le benjamin) est plus petit à la fois que le locuteur (l'aîné) et que le « premier » petit frère (le cadet ou puîné). Or il existe deux adjectifs pour exprimer 'petit' en romani : *tikno*, petit (qui serait utilisé pour le cadet), et *xurdo*, minuscule / tout petit (utilisé pour le benjamin). L'identification du référent se fait en opérant une restriction pragmatique, au sein des référents du nom (trois personnes, dont deux frères), à un seul référent, celui qui présente la propriété de l'adjectif. L'adjectif reçoit un focus étroit pour signaler cette propriété distinctive. Ce focus présuppose que l'énoncé serait faux si une autre entité du même ensemble avait été sélectionnée, par exemple le frère cadet.

(156) – (157) M24M, Llakatund, 12 août 2014

Cet échange est issu d'un entretien dirigé où le locuteur se présente.

1	*Ch: ə keći brešengo isinan tu o Bismir?*
2	*M24M: [rire]*
3	*Dnj: Kozom brešengo isinan?*
4	*M24M: Me. ə/ ə bišuštar.*
5	*Ch: ə/ Isi tu phrala, phenja?*
6	*M24M: Isi ma duj/ phenja ðe jek phral.*
7	*Ole isi Grećija, bešena.*
8	*Ch: Kaj Grećija?*

9	*M24M: ə/ Aθina.*
10	*Se/ mi phen i/ i tikni, bešela Germanija.*
11	*Ch: Isi but breša kaj bešena avrik?*
12	*M24M: ə/ ćerena/ pandž breš/ mo phral o tikno.*
13	*Eðe me phenja, o duj jek isi Germanija akana jek breš/ ćerela kaj đeli Germanija.*
14	*Se i/ i bari phen isi/ Grećija.*
15	*Isi pe maksemeza, pe romesa, ćerena buti othar.*
16	*Biɕinena šeja eðe othar.*
1	Ch : Euh, quel âge tu as Bismir ?
2	M24M : [rire]
3	Dnj : Quel âge tu as ?
4	M24M : Moi. Euh/ euh vingt-quatre.
5	Ch : Euh/ tu as des frères et soeurs?
6	M24M : J'ai deux/ sœurs et un frère.
7	Ils sont en Grèce, il y habitent.
8	Ch : Où ça en Grèce?
9	M24M : Euh/ à Athènes.
10	Parce que/ ma/ ma petite sœur, habite en Allemagne.
11	Ch : Ca fait longtemps qu'ils habitent à l'étranger ?
12	M24M : Euh/ ça fait/ cinq ans/ mon petit frère.
13	Et mes sœurs, les deux une est en Allemagne depuis maintenant un an/ ça fait qu'elle est partie en Allemagne.
14	Et la/ la grande sœur est/ en Grèce.
15	Elle est avec ses enfants, avec son mari, ils travaillent là-bas.
16	Ils vendent des habits là-bas aussi.

(156)	*Se/*	***m-i***	***phen***	***i/***	***i***	***tikn-i,***	*beš-ela*	*Germanija*
	parce que	POSS.1SG-F.SG	frère(F.SG)	ART.DEF.F.SG	ART.DEF.F.SG	petit-F.SG	habiter-3SG.PRES	Allemagne

« Et/ ma/ ma petite sœur, habite en Allemagne. »

Statut de l'information

Le référent « ma petite sœur » est accessible-inférable (*set relation*) puisque le locuteur a annoncé qu'il avait deux sœurs et un frère. Une information est cependant nouvelle : non l'existence du référent (du nom), mais son âge (l'information portée par l'adjectif). L'adjectif *petite* est donné lexicalement.

Structure de l'information

Le syntagme *mi phen i/ i tikni*, ma/ ma petite sœur, constitue le topique-propos de l'expression *bešela Germanija*, habite en Allemagne.

Interprétation sémantique

La partie nominale du syntagme, la sœur, fait partie du contexte, du présupposé, puisque le locuteur a déjà mentionné qu'il avait des frères et sœurs. Le contraste se fait donc *in presentia* en discours. Il s'agit donc de sélectionner une sœur parmi les deux, sélection qui se fait au moyen de l'adjectif, qui apporte un sème spécifique permettant d'identifier le référent. En d'autres termes, le référent de

l'expression globale est à l'intersection de l'ensemble {sœurs} et de l'ensemble {entité petites}. On ignore si l'autre sœur est, quant à elle, plus petite ou plus grande que le locuteur. L'identification du référent se fait en opérant une restriction pragmatique, au sein des référents du nom, à un seul référent, celui qui présente la propriété de l'adjectif. L'adjectif reçoit un focus étroit pour signaler cette propriété distinctive. Le focus présuppose que le fait d'habiter en Allemagne est faux pour une autre entité du même ensemble (la deuxième sœur) qui pourrait être sélectionnée. Notons que ce syntagme ne présente pas de rupture interne car l'hésitation a lieu non entre la partie « nominale » et la partie « adjectivale » du syntagme polydéfini, mais au sein de la partie adjectivale. Le locuteur hésite non sur le fait d'employer la polydéfinitude, mais sur la propriété distinctive à mentionner.

(157)	*ə/*	*ćer-ena/*	*pandž*	*breš*	***m-o***	***phral***	***o***	***tikn-o***.
	Euh	faire-3PL.PRES	cinq	année[M.PL]	POSS.1SG-M.SG	frère(M.SG)	ART.DEF.M.SG	petit-M.SG

« Euh/ ça fait/ cinq ans/ mon petit frère. »

Statut de l'information

Le référent « mon petit frère » est donné-inactif puisque le locuteur a annoncé qu'il avait deux sœurs et *un frère*. Une information est cependant nouvelle : non l'existence du référent (du nom), mais son âge (l'information portée par l'adjectif). Il me communique une propriété de ce frère, celle d'être plus jeune que lui. L'adjectif *petit* est nouveau lexicalement.

Structure de l'information

Le syntagme *mo phral o tikno*, mon petit frère, constitue le topique-propos de l'expression *ćerena/ pandž breš*, ça fait/ cinq ans.

Interprétation sémantique

Contrairement aux exemples précédents, l'identification du référent ne se fait pas en opérant une restriction pragmatique au sein des référents du nom, à l'intersection de l'ensemble {frères} et de l'ensemble {entités petites}, car le locuteur n'a qu'un seul frère. *Petit*, qui constitue le focus, n'est donc pas une propriété permettant de contraster le frère avec un autre frère, c'est-à-dire permettant l'identification. L'adjectif transmet une information certes, (le fait que le frère soit plus jeune que le locuteur), mais une information non-restrictive. On ignore d'ailleurs si le frère est plus jeune ou non que les sœurs : on apprend plus tard que l'une des deux sœurs est plus grande que le locuteur (donc que le frère), et qu'une autre est plus petite que le locuteur (mais on ignore toujours comment elle se situe par rapport au frère). On pourrait considérer que l'identification consiste à sélectionner le frère parmi l'ensemble {frères et sœurs}, c'est-à-dire un ensemble hypéronymique, plus large que celui du nom ou de l'adjectif. Le contraste se fait donc *in presentia* en discours.

(158) M71A, Korçë, 28 juillet 2014

L'exemple (158) est issu d'un conte dont les protagonistes principaux sont trois frères.[1] L'aîné est toujours nommé *o baro phral*, le grand frère ; le cadet n'est jamais désigné d'une manière particulière car ne réalise jamais d'action individuelle ; le benjamin est nommé *o geralo*, le chauve, ou bien *o geralo phral*, le frère chauve.

1	*M71A: Lja ta o lire.*
2	*Akova/ čhidja la jek dabata/ o geralo phral/ em našela/ džala pe phralende.*
3	*Vakjela "em phrala/ kaj isinen tumen/ vakjela/ kaj tumare love?"*
4	*"Po akova na"/ vakjerdas/ kaj/ xalas i džuvel lendar e love?*
5	*"Man dikh so love isi man"/ o keralo/ o geralo.*
6	XXX *keti buti kerdas o geralo/ o phrala lenge asćer.*
7	*Sar kerel love tusa muka?*
8	*"Me na džanava!"/ vakela okova/ "manđe ale kəsmeti"/ vakela/ "xale manđe love/ me džepate* XX.*"*
9	*"Aj/ ka našas amen"/ vakjela e phralenge.*
10	*Eh! Našena/ naš naš naš.*
11	*Sine jek manuš kaj bikinela godik.*
12	*Jek lafi/ jek lira/ sar vakjerena, jek loli.*
13	*Jek lafi/ jek lira.*
14	*Eh! Akova o baro raklo/ o/ o phral o baro/ mukhlas pe romnja deođende/ khamni/ džanes akana.*
1	M71A : Il prit les lires.
2	Celui-ci/ lui donna un coup/ le frère chauve (fou)/ et il partit/ il se rendit chez ses frères.
3	Il dit « Eh les frères/ où êtes-vous/ dit-il/ où est votre argent ? »
4	« Oui il non »/ dit-il/ où/ est-ce que la femme a dépensé leur argent ?
5	« Regarde-moi tout l'argent que j'ai moi »/ le chauve/ le chauve.
6	XXX combien le chauve travailla-t-il !/ leurs frères soldats.
7	Comment est-ce qu'il gagne de l'argent (?) ?
8	« Je ne sais pas moi ! »/ dit-il/ « il y a du monde qui est venu chez moi »/ dit-il/ « ils ont pris mon argent/ dans ma poche XX ».
9	« Et/ on va s'en aller, nous »/ dit-il à ses frères.
10	Eh ! Ils partent/ partent/ partent.
11	Il y avait un homme qui vendait de la pensée.
12	Un mot pour une lire/ comme on dit/ une pièce.
13	Un mot/ une lire.
14	Eh ! Ce grand garçon/ le/ le grand frère/ laissa sa femme (?)/ enceinte/ tu sais maintenant.

(158)	*Akova*	*o*	*bar-o*	*raklo/*	*o/*
	DEM.M.SG	ART.DEF.M.SG	grand-M.SG	garcon_non_rom(M.SG)	ART.DEF.M.SG
	o	***phral***	***o***	***bar-o/***	*mukh-l-as*
	ART.DEF.M.SG	frère(M.SG)	ART.DEF.M.SG	grand-M.SG	laisser-PRET-3.SG

[1] Le début du conte se trouve aux exemples (161) et (195), dont le corpus est transcrit et traduit p. 565.

p-e	*romn-ja*	*deodžende/*	*khamn-i/*	*džan-esa.*
POSS.REFL-F.SG.OBL	femme-F.SG.OBL	?	enceint-F.SG	savoir-2SG.PRES

« Ce grand garçon, le… le frère aîné laissa sa femme (?)/ enceinte, tu sais. »

Statut de l'information

Le référent « le grand frère » est donné-actif puisqu'il vient d'être mentionné dans le même énoncé (le référent de *o phral o baro*, le grand frère, est le même que celui de *akova o baro raklo*, ce grand garçon non-rom). Si le référent est donné-actif, le nom est donné-inactif lexicalement. L'adjectif est donné-actif lexicalement puisqu'il est présent dans le syntagme *akova o baro raklo*.

Structure de l'information

Le syntagme *o phral o baro*, le grand frère, constitue une apposition au syntagme *akova o baro raklo*, ce grand garçon. Ils sont tous deux le topique-propos de l'expression *mukhlas pe romnja deođende*, laissa sa femme enceinte.

Interprétation sémantique

Le contraste se fait *in presentia* en discours, puisque les trois frères ont été mentionnés auparavant en discours. La référence se calcule par l'adjonction des sèmes de l'adjectif et de ceux du nom : le référent de l'expression globale est à l'intersection de l'ensemble {frères} et de l'ensemble {entités grandes}. L'identification du référent se fait en opérant une restriction pragmatique au sein des référents du nom. L'adjectif apporte ici un sème spécifique qui sert à identifier le référent.

(159) F38A, Korçë, 4 août 2014

Cet échange est issu d'un entretien dirigé où la locutrice se présente et mentionne sa fille.

1	*Ch: Save čhibja džanesa/ romane?*
2	*F38A: Romane, šćip.*
3	*Ch: Grećišt džanes?*
4	*F38A: Na/ ni džanav.*
5	*Ch: Hmm/ te/ tire maksemja džana/ ti škola?*
6	*F38A: džana/ i čhaj o čhavo. Džana.*
7	*Ch: Isi jek školla akate?*
8	*F38A: Akəna ali i čhaj kato kampi/ i čhaj i vaver.*
9	*F16AČ: Hmm/ keći brešengi si ti bori?*
10	*F38A: Tikni isi/ dešušov/ dešupandž.*
1	Ch : Quelles langues tu parles, le romani ?
2	F38A : Romani, albanais.
3	Ch : Tu parles grec ?
4	F38A : Non/ je sais pas.
5	Ch : Hmm/ tes/ tes enfants ils vont/ à l'école ?

6	F38A : Il y vont/ le garçon la fille Ils y vont.
7	Ch : Il y a une école ici ?
8	F38A : La fille vient de rentrer du camp/ l'autre fille.
9	F16AČ : Hmm/ Quel âge a ta bru ?
10	F38A : Elle est petite/ quinze/ seize.

(159)	*Akəna*	*al-i*	*i*	*čhaj*	*kat-o*	*kampi/*
	maintenant	venir.PRET-3SG.F	ART.DEF.F.SG	fille(M.SG)	venant_de-ART.DEF.M.SG	camp(M.SG)
	i	***čhaj***	***i***	***vaver***.		
	ART.DEF.F.SG	fille(M.SG)	ART.DEF.F.SG	autre(F.SG)		

« La fille vient de rentrer du camp / l'autre fille. »

Statut de l'information

Le référent « l'autre fille » est donné-actif puisqu'il vient d'être mentionné dans la même phrase. Le nom *fille* est donné mais l'adjectif *autre* est nouveau lexicalement.

Structure de l'information

Le syntagme *i čhaj i vaver*, l'autre fille, constitue le topique-propos de *ali kato kampi*, est venue du camp.

Interprétation sémantique

La partie nominale du syntagme, la fille, fait partie du contexte, du présupposé. La locutrice a en effet annoncé qu'elle avait huit enfants, quatre filles et quatre garçons. L'une de ses filles est présente durant l'entretien, elle vient de mentionner un garçon et une fille qui vont à l'école. La fille dont il est question dans cette phrase, qui va tous les jours aux activités du « camp », est une autre que celle qui va à l'école. Il s'agit donc de l'une des deux autres, mais la mère ne précise pas laquelle. Pour sélectionner celle-ci, elle utilise l'adjectif *vaver*, autre. Le contraste se fait donc *in presentia* en discours. Le référent de l'expression globale est à l'intersection de l'ensemble {filles} et de l'ensemble {autres entités}. L'identification du référent se fait en opérant une restriction pragmatique, au sein des référents du nom grâce à l'adjectif. Notons que, comme dans l'exemple précédent, le syntagme polydéfini est juxtaposé afin d'apporter une précision – à ceci près qu'il n'est pas directement apposé à son antécédent, *i čhaj*, la fille, mais mis en phrase.

(160) M23M, Llakatund, 13 août 2014

L'exemple (160) est issu d'un entretien dirigé.

1	*Ch: ə/ đeljan ti škola?*
2	*M23M: Đelom/ barundom i škola ə mesme.*
3	*Na therdom mundsia te vazhdonav po opre se na therdam love.*
4	*O mundsie isine të pakta but tha na therdam mundsia.*

5	*Ch: ə/ savi profesja/ savo/ so ka oves? So kames te keres?*
6	*M23M: Važdonav sa i škola/ e arteki, važdonav sa opre/ i/ ti škola i bari.*
7	*Ch: Em so ka keres?*
8	*M23M: Te važdonav sa/ pər pjano.*
1	Ch : Euh/ Tu as été à l'école ?
2	M23M : J'y ai été/ j'ai terminé le lycée.
3	J'ai pas eu la possibilité de continuer plus haut parce qu'on avait pas d'argent.
4	Les possibilités étaient très limitées et on n'avait pas de possibilité.
5	Ch : euh/ quel métier/ quel/ qu'est-ce que tu vas faire ? qu'est-ce que tu voudrais faire ?
6	M23M : Je termine toute l'école/ des arts, je la fais jusqu'au bout/ la/ la grande école.
7	Ch : Et qu'est-ce que tu vas faire ?
8	M23M : Que je continue tout ça/ pour le piano.

(160)	*Važdon-av*	*sa*	*opre/*	*i/*	***t-i***	***škola***	***i***	***bar-i.***
	terminer-1SG.PRES	tout	haut	ART.DEF.F.SG	à-ART.DEF.F.SG	école(F.SG)	ART.DEF.F.SG	haut-F.SG

« Je la fais jusqu'au bout/ la/ la grande école. »

Statut de l'information

Le référent « la grande école » est donné-actif puisque le locuteur vient de déclarer vouloir faire l'école des arts, qui est certainement celle-ci. Si ce n'est pas le cas, il serait accessible-inférable (*set-relation*). Le nom *école* est donné et l'adjectif *grande* nouveau lexicalement.

Structure de l'information

Le syntagme *i škola i bari*, la grande école, constitue le topique-cadre de l'expression *na žanav*, je ne sais pas.

Interprétation sémantique

La partie nominale du syntagme, l'école, fait partie du contexte, du présupposé. Le locuteur a en effet déjà mentionné l'école des arts dans la phrase précédente. Le syntagme polydéfini « la grande école » vise à apporter une propriété supplémentaire à « l'école des arts ». La référence se calcule par l'adjonction des sèmes de l'adjectif et de ceux du nom. En d'autres termes, le référent de l'expression globale est à l'intersection de l'ensemble {école} et de l'ensemble {entités grandes}. En effet, « la grande école » s'oppose à « le lycée », littéralement « l'école moyenne », dont il vient d'être question. Le contraste se fait donc *in presentia* en discours. L'identification du référent se fait en opérant une restriction pragmatique, au sein des référents du nom grâce à l'adjectif.

(161) – (195) M46A, Korçë, 24 juillet 2014

L'extrait suivant constitue le début d'un conte, dont la suite est donnée dans avec l'exemple (158), dont le corpus se trouve p. 555.

1	*M46A: Me ka sikavav tće jek masali/ Aurora/ kaj si jek masali/ e trine phralengo.*
2	*Trin phrala Aurora/ lele pes em gele/ jek/ vaver themeste/ kaj te ćeren buti/ Australja.*
3	*Em/ o duj phrala Aurora/ o bare isine maj džangle.*
4	*O jek tikno phral isi/ vaćeren lesće o duj phrala bare/*
5	*Eh akowa isi delino.*
6	[rire]
	Ama on na isine delino/ kalođes na isi delino.
7	*Lelas/ ah?*
8	*Ch: O tikno?*
9	*M46A: O tikno.*
10	*ə/ lelas len ə/ o/ o gadžo.*
11	*Čhudas len ti buti e duje phralen ćerena buti.*
12	*Em lena love.*
13	*Lena šukar love.*
14	*E dilines mukhle les to ćher te čaravel e/ guruven.*
15	*O delino/*
16	[pause]
	Em akale duj phrala ćerena isine buti saro dive/ e gadžesar.
17	*O phral o tikno/ bešela isine/ e guruvensar.*
18	*Bešena isin/*
19	*Bešela isine kaj te čaravel e guruven.*
20	*Po isine jek gav/ barvalo/ šukar gav.*
21	*Em/ paš odole gaveste isine jek/ purt/ purt.*
22	*Em ə/ jek len/ bari len/ je/ ah, Drenica/ vaćerena isine oj.*
23	*Odole lenjate/ ə/ odole/ bare lenjate isine i kučedra/ dešuduje šerensar.*
1	M46A : Moi je vais t'apprendre un conte/ Aurore/ qui est un conte/ des trois frères.
2	Trois frères Aurore/ partirent et allèrent/ dans un/ autre pays/ pour travailler/ en Australie.
3	Et/ les deux frères Aurore/ les grands étaient plus expérimentés.
4	L'un le petit frère est/ ils lui disaient les deux grands frères/
5	Eh mais il est fou.
6	[rire]
	Mais lui il n'était pas fou/ (?) il n'est pas fou.
7	Il prit/ hein ?
8	Ch : Le petit ?
9	M46A : Le petit.
10	Euh/ le/ le paysan (litt. le Non-Rom) euh/ les embaucha.
11	Il les fit travailler les deux frères travaillèrent.
12	Et ils gagnèrent de l'argent.
13	Ils gagnèrent pas mal d'argent.
14	Le fou, ils le laissèrent à la maison pour faire paître les/ vaches.
15	Le fou/
16	[pause]
	Et ces deux frères travaillaient toute la journée/ avec le paysan (litt. le Non-Rom).
17	Le petit frère/ habitait/ avec les vaches.
18	Ils habitai/
19	Il habitait là pour faire paître les vaches.
20	Mais il y avait un village/ riche/ beau village.
21	Et/ près de ce village il y avait un/ pont/ pont.
22	Et euh/ une rivière/ grande rivière/ euh/ ah, Drenica/ elle s'appelait comme ça.
23	Dans cette rivière/ euh/ cette/ grande rivière il y avait le dragon/ à douze têtes.

(161)	*O*	*phral*	*o*	*tikn-o/*
	ART.DEF.M.SG	frère(M.SG)	ART.DEF.M.SG	petit-M.SG
	beš-ela	*isine/*	*e*	*guruv-en-sar.*
	habiter-3SG.PRES	être.3.PRET	ART.DEF.PL.OBL	vache-M.PL.OBL-INSTR

« Le petit frère/ habitait/ avec les vaches. »

Statut de l'information

Le référent « le petit frère » est donné-inactif car il vient d'être mentionné dans l'avant-dernière phrase, *e dilines* et *o delino*, le fou. Le benjamin a été désigné jusqu'à présent comme *o jek tikno phral*, 'l'un petit frère', puis *o delino*, le fou, l'imbécile. Lexicalement, le nom *phral*, frère, est donné-actif car il a été employé dans la phrase précédente, et l'adjectif *tikno*, petit, est donné-inactif car il a été cité en début de conte.

Structure de l'information

Le syntagme *o phral o tikno*, le petit frère, constitue le topique-propos de la phrase.

Interprétation sémantique

Le contraste se fait *in presentia* en discours, puisque les trois frères ont été mentionnés auparavant en discours. La référence se calcule par l'adjonction des sèmes de l'adjectif et de ceux du nom : le référent de l'expression globale est à l'intersection de l'ensemble {frères} et de l'ensemble {entités petites}. L'identification du référent se fait en opérant une restriction pragmatique au sein des référents du nom. L'adjectif apporte ici un sème spécifique qui sert à identifier le référent.

(195)	*Em/*	***o***	***duj***	***phral-a***	*Aurora/*	***o***
	et	ART.DEF.PL	deux	frère-M.PL	Aurore	ART.DEF.PL
	bar-e	*isine/*	*maj*	*džan-gl-e.*		
	grand-PL	être.3.PRET	COMP	savoir-PART.PASS-PL		

« Et/ les deux frères Aurore/ les grands étaient plus expérimentés. »

Le syntagme présente une insertion suivie d'une pause de 0,36 seconde.

Statut de l'information

Le référent est accessible inférable car il vient d'être question de « trois frères » dans la phrase précédente. Le nom est donné-actif, mais l'adjectif est nouveau lexicalement.

Structure de l'information

Le syntagme *o duj phrala Aurora/ o bare*, les deux grands/ frères Aurore, constitue le topique-propos de *isine maj džangle*, étaient plus expérimentés.

Interprétation sémantique

Le contraste se fait *in presentia* en discours, puisque les trois frères ont été mentionnés auparavant en discours. La partie nominale du syntagme, les deux frères, fait partie du présupposé. Le syntagme vise à sélectionner deux référents au sein de l'ensemble {frères}, grâce à l'adjectif *grands*. Le référent global est donc à l'intersection de ceux du nom et de ceux de l'adjectif. Il s'agit d'un syntagme polydéfini à effet de sélection classique… si ce n'est qu'il présente une rupture. Celle-ci se manifeste à la fois par un silence et par l'insertion d'un mot, ici un nom propre au vocatif, *Aurora*. L'adjectif n'en est pas moins un modifieur indirect.

(162) F16AČ, Korçë, 21 juillet 2014

L'exemple suivant est issu d'une conversation spontanée. Be et moi discutions des divers voyages qu'il a effectués en Europe. Sa fille, Br, entre dans la pièce au moment où il mentionne Istamboul. Entendant ce nom mais ignorant l'objet de la discussion, F16AČ s'exclame :

1	*F16AČ: Istanbul/ kaj ćerena o filme o bare!*
1	F16AČ : Istamboul, là où on fait les grands films !

(162)	*Istanbul/*	*kaj*	*ćer-ena*	***o***	***film-e***	***o***	***bar-e!***
	Istamboul	où	faire-3PL.PRES	ART.DEF.PL	film-PL	ART.DEF.PL	grand-PL
	« Istamboul, là où on fait les grands films ! »						

Statut de l'information

La remarque de F16AČ semble tomber du ciel et n'est pas liée à la conversation en cours. Dans la mesure où il n'est pas question de films (turcs ou non) dans la conversation, et où il n'en a pas été question lors des jours qui précèdent cette énonciation – pas en ma présence du moins –, on peut considérer le référent « les grands films » comme nouveau lexicalement et référentiellement. Il est cependant possible que du point de vue de la locutrice, il fasse partie de la culture générale et soit pour cette raison accessible-général. L'adjectif *grands* est nouveau lexicalement.

Structure de l'information

Le syntagme *o filme o bare*, les grands films, constitue le focus de la phrase, plus précisément un focus de nouvelle information non sollicitée.

Interprétation sémantique

On pourrait considérer que le référent soit à l'intersection de l'ensemble {films} et de l'ensemble {entités grandes}. Mais dans la mesure où il n'est pas question d'autres films, de moindre qualité ou venant d'autres lieux de production, il est difficile de parler de sélection d'un référent au sein d'un ensemble d'entités similaires. S'il y a contraste, il se fait *in absentia*, hors de la mention en discours, ou de la présence en situation, des entités avec lesquelles contraste le référent. En revanche, il fait bel et bien l'objet d'un focus qui souligne son sens : on peut parler d'un focus étroit à valeur assertive (ou informative, syntagmatique). La valeur emphatique de ce type de focus va ici de pair avec le fait que la phrase soit exclamative.

(163) M46A, Korçë, 15 août 2014

Cet exemple est issu d'une conversation spontanée. Il est prononcé par Be au cours d'un trajet en voiture dans Korçë. Alors que nous passons à côté d'un ensemble d'immeubles, d'un gros alambic de cuivre et d'un *Biergarten* attenant, il pointe tout à coup du doigt l'ensemble et s'exclame :

1	*M46A: Aćija si i bira i purani. Italjako vakət, to 1936. I proti bira ti Albanja.*
1	M46A : Ça c'est l'ancienne brasserie. À l'époque italienne, en 1936. La première brasserie en Albanie.

(163)	*Aćija*	*si*	***i***	***bira***	***i***	***puran-i.***
	DEM.F.SG	être.3.PRES	ART.DEF.F.SG	brasserie(F.SG)	ART.DEF.F.SG	vieux-F.SG
	« Ça c'est l'ancienne brasserie. »					

Statut de l'information

Le référent « l'ancienne brasserie » est accessible-situationnel car le référent fait physiquement partie de la situation d'énonciation. L'adjectif *ancienne* est nouveau lexicalement.

Structure de l'information

Le syntagme *i bira i purani*, l'ancienne brasserie, constitue le focus de la phrase, un focus de nouvelle information non sollicitée.

Interprétation sémantique

On pourrait considérer que le référent soit à l'intersection de l'ensemble {brasserie} et de l'ensemble {entités nouvelles}. Mais dans la mesure où l'ancienne brasserie est la seule de la ville, où la bière continue d'y être produite, il est difficile de parler d'un focus contrastif. Le locuteur n'oppose pas une ancienne usine à une nouvelle, mais veut me livrer une propriété de cette usine : elle est ancienne.

C'est pourquoi on peut plutôt parler d'un focus assertif. Il va de pair avec le fait que le référent soit annoncé par un déictique en début d'énoncé et physiquement montré du doigt. S'il y a contraste, il se fait *in absentia*, hors de la mention en discours, ou de la présence en situation, des entités avec lesquelles contraste le référent.

(164) F16AČ, Korçë, 24 juillet 2014

Au cours d'une promenade, F16AČ me montre du doigt un bâtiment et prononce ces mots :

1	*F16AČ: Me biandilom ando ćher odothe, kaj si akana o lokali o baro.*
1	F16AČ : Je suis née dans la maison là-bas, où il y a maintenant le grand café.

(164)	*Me*	*bian-dil-om*	*ando*	*ćher*	*odothe,*	*kaj*
	1.SG	accoucher-PASS.PRET-1SG	dans.ART.DEF.M.SG	maison(M.SG)	là-bas	où
	si	*akana*	***o***	***lokali***	***o***	***bar-o***.
	être.3.PRES	maintenant	ART.DEF.M.SG	café(M.SG)	ART.DEF.M.SG	grand-M.SG

« Je suis née dans la maison là-bas, où il y a maintenant le grand café. »

Statut de l'information

Le référent « le grand café » est accessible-situationnel car le référent fait physiquement partie de la situation d'énonciation. L'adjectif *grand* est nouveau lexicalement.

Structure de l'information

Le syntagme *o lokali o baro*, le grand café, constitue le focus de nouvelle information non sollicitée dans la proposition subordonnée relative, dont l'antécédent constitue le focus-cadre de la phrase.

Interprétation sémantique

On peut interpréter de deux manières le dessein de la locutrice :

- **Effet de sélection** : la locutrice vise à sélectionner un café parmi d'autres, celui qui est grand. Elle le montre donc du doigt en employant les déictiques *odothe*, là-bas, et *akana*, maintenant. Elle suppose que les autres cafés ne sont pas grands, ou sont moins grands que ceux auxquels réfère la structure polydéfinie. Les autres cafés sont présents en situation, donc le contraste se fait *in presentia*.

- **Effet d'emphase** : la locutrice me signifie implicitement que la maison dans laquelle elle est née et a passé les premières années de son enfance lui plaisait mieux que celle dans laquelle

elle vit au moment de l'énonciation – parce qu'elle était plus belle, plus grande (la perception des dimensions est différente chez un enfant et chez un adulte) et qu'elle y est en somme plus attachée. Elle met donc sous emphase la propriété *grand* du référent qu'elle désigne, c'est-à-dire, aujourd'hui, du café. Cette interprétation est rendue possible par le fait que tous les cafés de cette partie de la rue sont, selon moi, sensiblement de la même taille. Du moins cette propriété ne m'a-t-elle pas semblée (suffisamment) distinctive.

(165) – (166) F16AČ, Korçë, 24 juillet 2014

Les exemples suivants sont issus d'une conversation spontanée qui eut lieu quelques jours plus tard en feuilletant un album photo :

1	F16AČ: Akale isine sare o studentja e o/ miro dad ćerela isine buti/ ćerela. [pause]
2	Akate isi amaro ćher/ [bâillement]
3	Ch: Hmm.
4	F16AČ: kaj isi o lokal o baro/ ti/ katredale. [pause]
5	Othe isine amaro ćher o baro. Paše.
1	F16AČ : Eux ce sont tous les étudiants et le/ mon père travaillait/ il travaillait. [pause]
2	Ici c'est notre maison/ [bâillement]
3	Ch : Hmm.
4	F16AČ : Là où il y a le grand café/ à la/ cathédrale. [pause]
5	Là il y avait notre grande maison. À côté.

(165)	*Kaj*	*isi*	***o***	***lokal***	***o***	***bar-o/***	*ti/*	*katredale.*
	où	être.3.PRES	ART.DEF.M.SG	café(M.SG)	ART.DEF.M.SG	grand-M.SG	à.ART.DEF.F.SG	cathédrale(F.SG)

« Là où il y a le grand café/ à la/ cathédrale. »

Statut de l'information

Dans cet exemple, le référent est accessible-général car il fait partie de ma connaissance du monde, depuis la conversation de l'exemple (164), mais le nom et l'adjectif sont nouveaux lexicalement.

Structure de l'information

Le syntagme *o lokal o baro*, le grand/ café, constitue le topique-cadre de l'expression *amaro ćher*, notre maison.

Interprétation sémantique

On peut interpréter de deux manières le dessein de la locutrice :

- **Effet de sélection** : La partie nominale du syntagme, le café, fait partie du contexte, du présupposé, dans la connaissance du monde. On peut considérer que la locutrice sélectionne

un café parmi d'autres grâce à l'adjectif *grand* sous focus contrastif. Aucun des cafés n'est mentionné en discours ni n'est présent en situation : on peut dire que le contraste se fait *in absentia.*

- **Effet d'emphase** : Il se peut en effet que l'adjectif ne délivre une propriété du café mise sous emphase, donc sous focus assertif. La répétition de l'adjectif *baro,* grand, pour désigner à la fois la maison et le café, donne l'impression que la propriété d'être grand est passée de la maison d'enfance au café qui occupe à présent le rez-de-chaussée de l'immeuble. Il semble y avoir transmission de cette propriété de grandeur : si tel est le cas, il s'agit d'une propriété intrinsèque du référent désignée, en dehors de toute comparaison avec les cafés voisins.

(166)	*Othe*	*isine*	***amar-o***	***ćher***	***o***	***bar-o***.
	Là	être.3.PRET	POSS.1PL-M.SG	café(M.SG)	ART.DEF.M.SG	grand-M.SG

« Là il y avait notre grande maison. »

Statut de l'information

Le référent « le grand café » est accessible-général car le référent fait maintenant partie de ma connaissance du monde. L'adjectif *grand* est donné lexicalement, car il a déjà été employé dans la structure DEF N, DEF A (*o lokal o baro*) proche du DEF N DEF A (*o lokali o baro*) de l'exemple (164). Il semble que la propriété *grand* soit transférée du café à la maison : cela va dans le sens de la deuxième interprétation (focus étroit informatif) de l'exemple (164).

Structure de l'information

Le syntagme *amaro ćher o baro*, notre grande maison, constitue dans la phrase un focus de nouvelle information sollicitée.

Interprétation sémantique

On peut interpréter de deux manières le dessein de la locutrice :

- **Effet de sélection** : la locutrice vise à distinguer une maison, celle qui est grande, d'une autre, celle où elle vit actuellement. Elle oppose aussi les déictiques *akate*, ici, et *othe*, là. Cela suppose que l'autre maison (celle où elle vit actuellement, et où elle prononce alors cette phrase), ne présente pas la propriété d'être *grande*. Une telle restriction suppose une opposition binaire entre une maison actuelle, non-grande, et une maison ancienne, grande. L'ancienne maison est le lieu où a été prise la photo, et la nouvelle maison est celle où se tient la conversation : on peut dire que le contraste se fait *in presentia*.

- **Effet d'emphase** : la locutrice me signifie implicitement que la maison de sa prime enfance lui plaît mieux que celle dans laquelle elle a déménagé ensuite. Elle met donc sous emphase la propriété 'grand' de l'ancien référent, qui prend un sens de superlatif absolu (notre *très grande* maison) et non relatif (notre maison *la plus grande [des deux]*). Cette interprétation est rendue possible par le fait que, selon moi, la maison où elle vit actuellement n'est pas plus

petite que celle de sa prime enfance[1]. Du moins cette propriété ne m'a-t-elle pas semblé (suffisamment) distinctive.

(167) F66M, Voskopojë, 28 juillet 2014

L'extrait suivant est issu d'un entretien dirigé :

1	*F66M: Eh/ me ćaava avdije dešušov/ dešuduj butja.*
2	*Ćava jek aktiviteti andre me khereste ka sikavava vetəm e ćhiine tiknen ka na te džan*
3	*džungale.*
4	*Sikad ekh ə jekto klaso tha phirena but šukar.*
5	*Ćava buti e tiknensar/ ando/ gava/ ki Šiprija ka nane registrume ki džende civile.*
6	*Ćava buti/ parallel/ e romjensa e dijensa ka te Xnga len e tikne ki škola.*
7	*Ta te na/ maren len/ te na duden e tiknen/ por te sikaven len.*
8	*Me ćava buti e romensa ta/ tha e romjensa ose nane problemi o romnjako keć/*
9	*po but isi o roma tha o čhave o bare apo đeshta apo sastro apo sasuj.*
10	*Me ćava buti ko grupi tha thava suksesi dja akana/*
	thava but projekti/ XXXXX/ ko Soros si/ ko jek Amerikani/ ćaava but butja.
1	F66M : Eh/ moi aujourd'hui j'ai seize/ douze travaux.
2	Je travaille à l'intérieur/ chez moi/ j'enseigne seulement aux petites filles, à ce qu'elles ne deviennent pas sales.
3	J'ai enseigné en euh en CP et ça marche très bien.
4	Je travaille avec les petits/ dans les/ villages/ en Albanie là où ils ne sont pas inscrits au registre de l'état civil.
5	Je travaille/ en parallèle/ avec les femmes les mères pour qu'elles les XX envoient les petits à l'école.
6	Et pour que/ qu'elles ne les frappent pas/ qu'elles ne (?) pas les petits/ mais qu'elles leur montrent les choses.
7	Moi je travaille avec les hommes pour que/ et les femmes/ et le problème, le mal ne vient pas des femmes/
8	mais bien plutôt des hommes et des grands enfants ou du beau-frère? ou du beau-père ou de la belle-mère.
9	Moi je travaille avec des groupes/ et ça a bien marché jusqu'à aujourd'hui/
10	j'ai beaucoup de projets/ XXXXX/ c'est chez Soros/ chez un Américain/ j'ai beaucoup de travaux.

(167)	*Ose*	*nane*	*problemi*	*o*	*romnj-a-k-o*	
	ou	être.NEG.3.PRES	problème	ART.DEF.M.SG	femme_Rom-OBL.SG-GEN-M.SG	
	keq/	*po*	*but*	*isi*	*o*	*rom-a*
	mal(M.SG)	plus	beaucoup	être.3.PRES	ART.DEF.PL	homme_Rom-M.PL

[1] C'est même l'inverse qui m'a semblé vrai : la maison actuelle est plus spacieuse que l'ancienne maison.

tha	***o***	***čhav-e***	***o***	***bar-e***.
et	ART.DEF.PL	enfant-M.PL	ART.DEF.PL	grand-PL

« Et le problème ne vient pas des femmes/ mais bien plutôt des hommes et des grands enfants […]. »

Statut de l'information

Le référent « les grands enfants » est accessible-inférable (*accessible-inferable set-relationship* si l'on suit la typologie établie par Götze et al. (2007). En effet, il est question d'enfants à plusieurs reprises dans la conversation : *e ćhiine tiknen*, les petites filles, *e tikne, e tiknen*, les petits. On pourrait aussi arguer que le référent est accessible-général puisqu'il fait partie de la culture générale que les femmes mariées peuvent avoir des enfants. L'adjectif *grands* est nouveau lexicalement.

Structure de l'information

Le syntagme *o čhave o bare*, les grands enfants, constitue avec *o Roma*, les adultes, un focus contrastif de remplacement, *contrastive subtype replacing* (Götze et al. 2007).

Interprétation sémantique

Dans cet exemple, deux interprétations sont possibles :

- **Effet de sélection** : parmi l'ensemble des enfants possibles, la locutrice sélectionne les enfants *grands* comme référents. Elle restreint le choix des référents à ceux qui sont à l'intersection de {enfants} et de {entités grandes}, excluant ceux à l'intersection de {entités petites}. Ceux qui posent problème seraient donc explicitement les grands et non les petits enfants, considérés comme inoffensifs (*te na duden e tiknen*, qu'elles ne (?) pas les petits). Comme les femmes adultes, *o romnja*, et les hommes, *o roma*, viennent d'être mentionnés, on peut dire que le contraste se fait *in presentia*.

- **Effet d'emphase** : la locutrice entend signifier que c'est le fait d'être grand qui pose problème – d'où la mention des adultes (*o Roma*, les hommes adultes) et des adolescents (*o čhave o bare*, les grands enfants). Dans cette interprétation, la locutrice n'établit pas tant un *contraste* entre enfants petits et grands qu'un *parallèle* entre hommes adultes et enfants grands/adolescents, donc quasi-adultes. On pourrait aller jusqu'à dire que ce qui pose problème, c'est le fait qu'ils soient grands, donc aient des besoins et des exigences d'adultes. Dans cette interprétation, l'adjectif mis sous emphase revêtirait une valeur presque causale.

  Une deuxième interprétation du focus informatif consiste à dire que ce qui pose problème, c'est qu'ils soient de genre masculin, puisqu'elle dit explicitement que les femmes ne posent pas problème. En effet, *čhave* est ambigu et peut désigner soit les enfants en général, soit les fils (enfants mâles) en particulier.

(168) – (186) M68A, Voskopojë, 15 juillet 2013

L'exemple suivant est issu d'un entretien dirigé, mené à Voskopojë lors du congrès pour les droits des minorités :

1	*Ch: Savi čhib ə/ vaćeres?*
2	*M68A: Me miri čhip isi/ kabudžija.*
3	*Ch: ərlija?*
4	*M68A: Erlija/ sar phenen atə/ men/ me phenasa kabudžija.*
5	*Ame sinam ə/ štar fisja bare/ ti Albanja.*
6	*Kaj bučhona pes/ Kabudžija/ <Mečkar/>*
7	*Ch: <So isi o Kabudžija?>*
8	*M68A: čergar/ Kurtofja.*
9	*Ch: Kurtofja, so si Kurtofja?*
10	*M68A: Kurtofja si/ o Roma/ kaj pherena daha ə/ karvanensa.*
11	*Karəna pes/ maškare lende/ pire gada romane.*
12	*Kaj si len odja godik ə/ daha/ romani godik traditja/ kultura/ so la me.*
13	*Me kaj phenava/ isinam roma/ kaj/ džasa/ pal/ to ćidipe to/ gadžikane.*
14	*Našadam. Nanaj/ našadam. Po.*
15	*Akade o roma ə/ Kurtofja isi/ but paše ki traditja e romengi.*
1	Ch : Quelle langue euh/ tu parles ?
2	M68A : Moi ma langue c'est le kabudji.
3	Ch : Arli?
4	M68A : Arli/ comme ils disent/ nous/ moi on dit kabudji.
5	Nous on a euh/ quatre grands groupes/ en Albanie.
6	Qui s'appellent/ Kabudji/ <Mečkar/>
7	Ch : <Qu'est-ce que c'est, les Kabudjis ?>
8	M68A : Čergar/ Kurtofi.
9	Ch : Kurtofi, qu'est-ce que c'est que Kurtofi?
10	M68A : Les Kurtofis ce sont/ des Roms/ qui se déplacent encore euh/ en caravanes.
11	Ils se font eux-mêmes/ entre eux/ leurs habits roms.
12	Qui ont cet esprit euh/ encore/ cette pensée romani la mentalité la tradition/ culture/ je crois.
13	Moi ce que j'en dis c'est que/ nous on est des Roms/ qui/ on va/ on participe/ à des réunions/ non-roms.
14	On l'a perdue. C'est pas/ on l'a perdue. Oui.
15	Ces Roms-là euh/ les Kurtofis ils sont/ très proches de la tradition des Roms.

(168)	*Isin-am*	*Rom-a/*	*kaj/*	*dž-asa/*	*pal/*
	être-1PL.PRES	Rom-M.PL	REL	aller-1PL	après
	t-o	***ćidipe***	***t-o/***	***gadžikan-e.***	
	à-ART.DEF.PL	rassemblement[M.PL]	à-ART.DEF.PL	non_rom-PL	

« Nous on est des Roms/ qui/ on va/ on participe/ aux réunions/ non-roms. »

Statut de l'information

Le référent « les réunions non-roms » est accessible-situationnel car le référent fait partie de la situation d'énonciation : nous nous trouvons précisément à un rasssemblement organisé par des Non-Roms. L'adjectif *non-rom* est nouveau lexicalement.

Structure de l'information

Le syntagme *to ćidipe to/ gadžikane*, aux réunions/ non-roms, constitue avec le verbe le focus de la phrase (focus de nouvelle information non sollicitée).

Interprétation sémantique

- **Effet de sélection** : parmi l'ensemble des réunions auxquelles peut participer un groupe de Roms, le locuteur sélectionne les réunions des Non-Roms. Il vise à contraster de manière binaire deux profils de réunions : celles organisée par les Non-Roms, à leur manière et dans leur tradition, et celles organisée par les Roms, à leur manière et dans leur tradition – ce qui n'empêche d'ailleurs pas le public d'être mixte. S'il y a contraste, c'est *in absentia*, le locuteur présupposant que je sais qu'il existe des réunions non-roms, en quoi elles consistent et en quoi elles diffèrent des réunions roms. Les vêtements non-roms n'ayant pas été mentionnés, on peut dire que le contraste se fait *in absentia*.
- **Effet d'emphase** : au sein d'une conception anthropologique opposant les Roms au reste des ethnies, le locuteur entend signifier l'ouverture dont fait preuve son groupe, puisqu'il participe à des réunions organisées par les autres ethnies. Comme il n'est nullement question de *ćidipe romane*, réunions roms *stricto sensu*, on peut penser que ce n'est donc pas tant l'opposition binaire réunion rom ~ non-rom qui l'intéresse que l'idée d'ouverture culturelle – ou de dilution possible de l'identité (selon la perspective adoptée).

(186)	*Kar-əna*	*pes/*	*maškare*	*len-de/*	***pir-e***	***gad-a***	***roman-e***.
	faire-3PL.PRES	3.REFL	entre	3.PL.OBL-DAT	POSS.3REFL-PL	vêtement-M.PL	rom-PL

« Ils se font eux-mêmes/ entre eux/ leurs habits roms. »

Statut de l'information

Le référent est nouveau. Le nom est nouveau lexicalement et l'adjectif est nouveau lexicalement – il pourrait être accessible-inférable à cause de la mention du mot Rom à la phrase précédente.

Structure de l'information

Le syntagme *pire gada romane*, leurs habits à la façon rom, constitue le focus de nouvelle information non sollicitée, avec le verbe *karəna pes*, ils se font.

Interprétation sémantique

La référence se calcule par l'adjonction des sèmes de l'adjectif et de ceux du nom : le référent de l'expression globale est à l'intersection de l'ensemble {vêtements} et de l'ensemble {entités roms}. Il existe donc un focus étroit à valeur contrastive portant sur l'adjectif, qui s'oppose à l'adjectif *gadžikane*, non-roms. Le syntagme en (186) est un exemple de modification indirecte. Les réunions roms n'ayant pas été mentionnées, on peut dire que le contraste se fait *in absentia*.

(169) F24A, Bilisht, 27 juillet 2014

Mon hôte a réalisé un reportage télévisé sur les conditions de vie misérables des Roms dans les bidonvilles, ici à Bilisht. J'ai ensuite interviewé une femme vivant dans un bidonville.

1	*F24A: Kaj te vaćerav tuće/ kan aalo o Beni akate/ amen isinam isine fare/ katastrof.*
2	*Amen na isine amen maro/ te xasa isine.*
3	*Aale lele amen to iv/ jo ka anas tumenđe ndihme, jo okole tumenđe ndihme.*
4	*Amen čam avil o vend/ amen nanaj amen šeja te urjaras men.*
5	*Ka pere to drom vendesko em ka aves/ ka dikhes.*
6	*Amare kuške so čhidas amen/ hi li o šeja.*
7	*O xabe si o problemi o baro.*
8	*Po na xalas o manuš/ merela.*
9	*Je na xala jek dive.*
10	*Ka koxaves o per jek dive.*
11	*Lačho/ me ka durinav.*
12	*Por akova maksemi/ durinela?*
1	F24A : Il faut que je te dise/ quand Beni est venu ici/ nous on avait rien/ la catastrophe.
2	On n'avait pas une miette de pain/ pas de quoi manger.
3	Ils sont venus ils nous ont mis dehors dans le froid/ oui on va vous apporter de l'aide, oui ci et ça pour vous, de l'aide.
4	Pour nous dès que vient l'hiver/ nous on a pas de vêtements pour s'habiller.
5	Tu vas (?) sur le chemin de l'hiver et tu vas venir/ tu verras.
6	Nos (?) qu'on ramasse/ même nos vêtements.
7	La nourriture c'est le grand problème.
8	Oui si l'homme ne mange pas/ il meurt.
9	On ne mange pas pendant un jour entier.
10	Tu peux tromper ton ventre pendant un jour.
11	Bien/ moi je peux supporter ça.
12	Mais ce bébé-là/ est-ce qu'il supporte ?

(169)	*O*	*xabe*	*si*	***o***	***problemi***	***o***	***bar-o.***
	ART.DEF.M.SG	nourriture(M.SG)	être.3.PRES	ART.DEF.M.SG	problème(M.SG)	ART.DEF.M.SG	grand-M.SG

« La nourriture c'est le grand problème. »

Statut de l'information

Le référent « le grand problème » est accessible-inférable car la locutrice a déjà expliqué dans le discours qui précède en quoi la nourriture constitue un problème. L'adjectif *grand* est nouveau lexicalement.

Structure de l'information

Le syntagme *o problemi o baro*, le grand problème, constitue le topique-propos de l'expression *o xabe*, la nourriture.

Interprétation sémantique

- **Effet de sélection** : parmi l'ensemble des problèmes possibles, la locutrice sélectionne la nourriture. Elle vient d'expliquer d'une manière extensive, dans les minutes qui précèdent, toutes les difficultés économiques et sociales qu'elle rencontre pour survivre et élever ses enfants. Parmi l'ensemble {problèmes}, elle sélectionne la nourriture. Dans les quelques phrases qui précèdent, elle a évoqué le froid de l'hiver, le manque de vêtement, l'absence d'électricité : le contraste se fait donc *in presentia*. La structure DEF N DEF A confère à l'adjectif un sens superlatif relatif : on pourrait traduire *baro* par *principal, majeur*.
- **Effet d'emphase** : La structure DEF N DEF A pemet une sorte de bilan. La locutrice synthétise tout son discours en élisant comme problème majeur la nourriture. Dans cette interprétation, l'adjectif mis sous emphase donne l'impression que la nourriture est le seul problème qui l'inquiète : On pourrait traduire baro par *important*.

2.5. Corpus spontané : DEF A DEF N

(261) M71A, Korçë, 28 juillet 2014

Cette conversation spontanée se déroule chez mon hôte, dans la cour de la maison, après un entretien dirigé.

1	*M71A: Ame xatarda ames akana, eftavar' breš.*
2	*Kerava me šov/ pendupanč breš kaj pijava cigara.*
3	*Pendupanč/ po te gines en/ duj khera love si!*
	[rire]
4	*O love ma te astares! Duj khera love!*
	[longue pause]
5	*Ch: Ti romni kaj isi?*
6	*M71A: I romni bešli pe čat.*
7	*Akate si o džamutro, Korča.*

8	*"Beš, beš" na mukhle la o čhaja, o tikne o maksemi.*
	[pause]
9	*I on te pašljovav da temin othe.*
10	*Ch: Hmm*
11	*M71A: Isi man jek nuka/ pa oj/ merela mange.*
12	*Bracerdom la tikni.*
13	*"Papu papu, haj te čhuvav te pire."*
14	*"Na X mi chaj.*
15	*"Ka th'av man t-i česma."*
16	*"Na".*
17	*Ali thodas me pire.*
18	*But ma oj, but manđela man.*
1	M71A : Nous on pourrit maintenant, 70 ans.
2	Moi ça fait cinquante-ci/ cinquante-cinq ans que je fume des cigarettes.
3	55/ oui si tu comptes/ ça fait deux maisons en argent !
	[rire]
4	L'argent tu ne peux pas le retenir ! Deux maisons en argent !
	[longue pause]
5	Ch : Elle est où ta femme ?
6	M71A : La femme elle est restée chez (?).
7	Ici il y a le beau-fils/ à Korçë.
8	"Reste, reste", les filles, les petits enfants ne l'ont pas laissée [partir].
	[pause]
9	(?)
10	Ch : Hmm.
11	M71A : J'ai une petite-fille, elle/ elle m'aime à mourir.
12	Je l'ai bercée toute petite.
13	"Papi, papi, viens que je lave tes pieds."
14	"Non X ma fille."
15	"Je vais me laver à la fontaine."
16	"Non."
17	Mais elle a lave mes pieds quand même.
18	Elle me beaucoup [sic], elle m'aime beaucoup.

(261)	*Na*	*mukh-l-e*	*la*	*o*	*čhaj-a/*	***o***	***tikn-e***	***o***	***maksemi.***
	NEG	laisser-PRET-PL	3.SG.F.OBL	ART.DEF.PL	fille-F.PL	ART.DEF.PL	petit-PL	ART.DEF.PL	enfant(M.SG)

« Les filles, les petits enfants ne l'ont pas laissée [partir]. »

Statut de l'information

Le référent « les petits enfants » est donné-actif puisqu'il est le même que celui du syntagme « les filles » auquel il est apposé. Le nom *maksemi*, enfants, et l'adjectif *tikne*, petits, sont nouveaux lexicalement.

Structure de l'information

Le syntagme *o tikne o maksemi*, les petits enfants, constitue avec *o čhaja*, les filles, le topique-propos de l'expression *na mukhle la*, ne l'ont pas laissée.

Interprétation sémantique

La référence se calcule par l'adjonction des sèmes de l'adjectif et de ceux du nom : le référent de l'expression globale est à l'intersection de l'ensemble {enfants} et de l'ensemble {entités petites}. L'identification du référent se fait en opérant une restriction pragmatique au sein des référents du nom. L'adjectif apporte ici un sème spécifique qui sert à identifier le référent d'une manière que le locuteur estime plus précise que le syntagme sans modifieur *o čhaja*, les filles, auquel il est apposé. En effet, cette dernière expression est ambiguë : elle peut désigner aussi bien les filles que les petites-filles, donc deux générations différentes. *Petits enfants* fonctionne comme un concept unique, un nom composé avec une référence globale, d'où l'antéposition de l'adjectif.

(171) – (193) [Zamira][1], Korçë, 1er août 2014

Cette conversation spontanée se déroule après un passage au cimetière visant à rendre hommage aux morts. En cette occasion, la famille achète aux défunts de quoi boire, manger et fumer – ce qui représente une dépense pécuniaire. Au milieu du brouhaha général, la locutrice s'adresse à sa belle-sœur (la sœur de son mari), en reprochant au frère de celle-ci[2] (son beau-frère) de n'avoir pas participé aux dépenses.

1	*Zamira: Isine to gròbja.*
2	*Ch: Hmm.*
3	*Zamira: Pe dadesće.*
	[plusieurs voix]
4	*Zamira: Sora kale/ čhave/ sare akale nudja.*
5	*Beš ke X!*
6	*Kon čhudja XXX Fatma !*
	[brouhaha]
7	*Zamira: He? O barvalo o Skender so kerdas? Pesko làfi? Ničumuni!*
	[plusieurs voix]
8	*Zamira: So kerdas to/ to to phral/ o barvalo?*
9	*XXX ti kangeri/ aj te XXX?*
	[plusieurs voix]
10	*Fatma: te čhuves anre.*
11	*Zamira: Eh/ čhuv tu kaj isinan barvalo!*
12	*Kaj ka čhuvav dujšel trinšel xiljade lek kaj si odole dyšind trešind mij lek!*
	[plusieurs voix]
13	*Zamira: Na gelan to phure č Aurora? To limòra na gelan? To mejdùšan, to gròbja?*
1	Zamira : C'était au cimetière.
2	Ch : Hmm.
3	Zamira : Pour le père.
	[plusieurs voix]
4	Zamira : Ensuite ces/ enfants/ tous ses petits enfants/

[1] La locutrice a été anonymisée.

[2] Les prénoms de toutes les locutrices et de tous les protagonistes ont été modifiés afin de préserver leur anonymat.

5	Attends que X !
6	Qui a mis XXX ? Fatma !
	[brouhaha]
7	Zamira : Eh ? Skender le riche qu'est-ce qu'il a fait ? sa parole ? rien !
	[plusieurs voix]
8	Zamira : Qu'est-ce qu'il a fait ton/ ton ton frère/ le riche ?
9	XXX à l'église/ et que XXX ?
	[plusieurs voix]
10	Fatma : pour que tu mettes les oeufs
11	Zamira : C'est ça/ mets-les/ toi qui es riche !
12	Et c'est moi qui vais y mettre 200, 300 mille$_{romani/grec}$ lires, où sont ces 200, 300 mille$_{albanais}$ lires !
	[plusieurs voix]
13	Zamira : T'as pas été chez les vieux, eh Aurore ? T'as pas été au cimetière$_{grec}$? Au cimetière$_{turc}$, au cimetière$_{slave}$?

(171) ***O*** ***barval-o*** ***o*** ***Skender*** *so* *ker-d-as?*

ART.DEF.M.SG riche-M.SG ART.DEF.M.SG Skender(M.SG) INT faire-PRET-3SG

« Skender le riche qu'est-ce qu'il a fait ? »

Statut de l'information

Le référent « Skender le riche » est accessible-général car c'est un membre de la famille connu de la locutrice comme de l'interlocutrice. L''adjectif *barvalo*, riche, est nouveau lexicalement.

Structure de l'information

Le syntagme *o barvalo o Skender*, Skender le riche, constitue constitue le topique-propos de l'expression *so kerdas*, qu'est-ce qu'il a fait.

Interprétation sémantique

Là aussi, l'identification du référent se fait en opérant une restriction pragmatique. Deux possibilités d'interprétation s'offrent à nous : le référent de l'expression globale peut être à l'intersection de l'ensemble de {personnes nommées Skender} et de {personnes riches}, ou bien à l'intersection de {personnes de la famille} et de l'ensemble {personnes riches}. Dans le premier cas, elle présuppose l'existence de plusieurs Skender, parmi lesquels elle en sélectionne un, celui qui est riche. Dans le deuxième, elle présuppose l'existence d'une famille, au sein de laquelle elle sélectionne un membre qui présente la propriété d'être riche. Nous penchons pour la deuxième interprétation. En effet, il n'existe aucun autre Skender dans les environs. La locutrice reproche à son beau-frère de n'avoir pas participé aux dépenses pour les morts de sa famille. Il aurait dû d'autant plus participer à ces dépenses qu'elle le considère plus riche qu'elle. On a donc un contraste sous-jacent entre *ton frère*, riche, et le reste de la famille, moins riche. Plusieurs combinaisons sont d'ailleurs possibles pour le reste de la famille : {moi}, {moi + mon mari}, {moi + toi} ou {moi + toi + mon mari}.

(193)	*So*	*ker-d-as*	*t-o/*	***t-o***	***phral/***	***o***	***barval-o****?*
	quoi	faire-PRET-3SG	POSS.2SG-M.SG	POSS.2SG-M.SG	frère(M.SG)	ART.DEF.M.SG	riche-M.SG

« Qu'est-ce qu'il a fait ton/ ton ton frère/ le riche ? »

Le syntagme présente une pause de 0,27 seconde.

Statut de l'information

Dans cet exemple, le référent est donné-actif puisqu'il vient d'être mentionné dans la phrase précédente. Mais comme l'interlocutrice a fait la sourde oreille à la phrase précédente, le référent est censé être nouveau pour elle. Le nom *frère* est nouveau, mais l'adjectif *riche* est donné-actif lexicalement – ou nouveau, selon la perspective qu'on adopte.

Structure de l'information

Le syntagme *to phral/ o barvalo*, ton frère/ le riche, constitue le topique-propos de l'expression *so kerdas*, qu'est-ce qu'il a fait.

Interprétation sémantique

La partie nominale du syntagme, frère, fait partie du contexte, du présupposé, dans la connaissance du monde. L'adjectif peut viser à sélectionner un frère parmi les autres grâce à l'adjectif *riche* sous focus contrastif : au sein de l'ensemble {frère}, elle choisit le référent Skender (son beau-frère) et non le référent Enver[1] (son mari) car le premier est riche et le second non. Mais on peut aussi considérer que l'adjectif est mis sous emphase dans le but de critiquer le référent, grâce à un focus assertif. Le contraste se fait *in absentia*.

2.6. Corpus spontané : DEF A N

(172) F60A, Korçë, 21 juillet 2014

Cet échange est issu d'un entretien dirigé, où la locutrice présente les langues qu'elle parle.

1	*F60A: Me ćerava dialekti/ erlijasko/ er/ erlija.*
2	*M46A: Tu a/ a/ džanesa e Čergarengi čhib ose/ aćharesa la/ <des la?>*
3	*F60A: <Me> kan ćerava lensa lafi me aćarava i Čergarengi čhib.*
4	*M46A: Tu/ sostar?*
5	*F60A: E/ sostar se mə/ otkha otkha jek isi amari čhib ə/ sa i romani čhib jek isi.*

[1] Le prénom a été modifié.

6	*<Hać̦arava la.>*
7	*M46A: <XXX>*
8	*Les tut kan <lensa.>*
9	*F60A: <Po.>*
10	*I li goslavjać̦e Romensar pelo/ manđe jek dive to treni/ arakhlam amen em ə/ ać̦arava lengi čhib isine.*
1	F60A : Moi je parle le dialecte/ erlija/ er/ erlija.
2	M46A : Toi a/ a/ tu parles la langue des Čergars ou/ tu la comprends/ <tu la parles ?>
3	F60A : <Moi> quand je parles avec eux moi je comprends la langue des Čergars.
4	M46A : Toi/ pourquoi ?
5	F60A : Eh/ pourquoi parce que moi/ et comme ça notre langue est unitaire euh/ toute la langue romani est une et une seule langue.
6	<Je la comprends.>
7	M46A : <XXX>
8	Tu les entends <avec eux.>
9	F60A : <Oui.>
10	Même les Roms yougoslaves un jour je suis tombée/ sur eux dans le train/ on s'est rencontrés et euh/ je comprenais leur langue.

(172)	***Sa***	***i***	***roman-i***	***čhib***	*jek*	*isi.*
	tout	ART.DEF.F.SG	romani-F.SG	langue(F.SG)	IND	être.3.PRES

« Toute la langue romani est une et une seule. »

Statut de l'information

Dans cette conversation, le référent « langue romani » est mentionné à deux reprises : *amari čhib*, notre langue, puis *sa i romani čhib*, toute la langue romani. La première occurrence est nouvelle – mais l'on pourrait en réalité considérer qu'elle est accessible-inférable (*accessible-inferable set-relationship*, Götze et al. 2007) car le référent « langue romani » est cognitivement préparé par la mention de *dialekti/ erlijasko*, dialecte/ de l'arli, et *de e čergarengi čhib*, la langue des Čergars. La deuxième occurrence, qui nous intéresse ici, est donnée-active puisque le référent vient d'être cité dans la même phrase. Plus précisément, le syntagme nominal est donné référentiellement, le nom donné lexicalement, mais l'adjectif nouveau lexicalement.

Structure de l'information

Le syntagme *sa i romani čhib*, toute la langue romani, constitue le topique-propos (*aboutness topic*) de l'expression *jek isi*, est une et une seule.

Interprétation sémantique

La référence se calcule par l'adjonction des sèmes de l'adjectif et de ceux du nom : le référent de l'expression globale est à l'intersection de l'ensemble {langue} et de l'ensemble {entités romani}. On peut cependant douter que, dans ce contexte, l'identification du référent se fasse en opérant une restriction pragmatique au sein des référents du nom puisqu'il n'est question d'aucune autre langue que de la langue romani. En outre l'adjectif a ici une valeur globale, contrairement aux termes précédents qui désignent telle ou telle variété particulière (arli et čergar) – à l'instar également du

nom *langue*, d'extension plus grande que le nom *dialecte*. Ceci est confirmé par le déterminant *sa*, toute, et par la prédication *jek isi*, est une. L'adjectif *romani* est donc moins spécifique que *erlija*, *čergarçe*, et se place donc avant le nom : *romani čhib* apparaît comme une forme d'hypéronyme. Le syntagme en (172) est donc un exemple de modification directe.

(173) M71A, Korçë, 28 juillet 2014

Cet échange est issu d'un entretien semi-dirigé, où le locuteur raconte sa vie en Grèce.

1	*Ch: Nanaj/ nanaj grobja/ musulman?*
2	*M71A: Nanaj grobja musliman.*
3	*Sa gadženge.*
4	*Ch: Hmm.*
5	*M71A: Gadžikane grobja.*
6	*Ch: Nanaj sar akate/ Korča.*
7	*M71A: Na.*
8	*Ch: Kaj isi but.*
9	*M71A: Akate si koraxake/ si/ romanenge si/ čhibane.*
10	*Čhibane Roma vaćerasa lenge amen.*
11	*Ch: Na Eđipqan?*
12	*M71A: Eđyptjan/ ka vaćeras ama na/ vaćasa ole čhibane Roma.*
13	*Ch: Ahah.*
14	*M71A: Kaljarde si s/*
15	*Ch: Va.*
16	*M71A: Akana si parne ole po/ vaćerasa amen čhibane Roma amen akal ačhilo.*
17	*Ch: On si Roma i li/ na džanena i čhib/ <roma/>*
18	*M71A: <Na džane>na i čhib/ čhib.*
19	*Ch: Sosće na džanena <i čhib?>*
20	*M71A: <Na džane>na i čhib.*
21	*S/ čhibane Roma vaćerasa <lenge>.*
22	*Ch: <Sosće> na džanena <i čhib?>*
23	*M71A: Na džanena sće naj Roma!*
24	*Ch: Nanaj <Roma?>*
25	*M71A: <As>ja!*
26	*Nanaj asi Roma.*
27	*Ch: Hmm.*
	[silence]
28	*Sea ki lenge Eđyptjanenge.*
29	*Ch: Katar i Eđypt/ <Eđypta aven.>*
30	*M71A: <Eđypt/ Eh.>*
31	*Ch: Kana avile on? Džanesa tu?*
32	*M71A: Eh akowa phirade otakal/ to purane breša.*
33	*O Eđypti nakhle o mileti ondili čingar.*
34	*Phirade o mile.*
35	*Hanga odo nakhle sar/ amen e Albanija/ kan nashle o mileti/ Đermanija Italija Franca/ Anglija.*

36	*Kaj na đele!*
37	*Em phiradile akana!*
38	*Vaćasa albaneze.*
39	*E a Grećija si/ trin gava kaj džanena čhibane.*
40	*Amari čhib ćerena/ albanika.*
41	*Ch: Hmm.*
42	*M71A: Katar si ole? Albanijatar.*
43	*Ondile Greća.*
1	Ch : Il n'y a pas/ il n'y a pas de tombes/ musulmanes?
2	M71A : Il n'y a pas de tombes musulmanes.
3	Toutes de chrétiens [litt. de Non-Roms].
4	Ch : Hmm.
5	M71A : Des tombes chrétiennes.
6	Ch : C'est pas comme ici/ à Korçë.
7	M71A : Non.
8	Ch : Où il y en a beaucoup.
9	M71A : Ici il y en a de Turcs/ il y en a/ de Roms il y en a/ égyptiennes [albanaises].
10	Roms albanais [litt. bavards], on les appelle comme ça nous.
11	Ch : Pas *Égyptiens* ?
12	M71A : Égyptiens/ on dit mais non/ on les appelle Roms albanais.
13	Ch : Ahah.
14	M71A : Ils sont basanés s/
15	Ch : Oui.
16	M71A : Maintenant ceux-là ils sont blancs mais/ nous on dit appelle Roms albanais nous, c'est resté comme ça.
17	Ch : Eux c'est des Roms ou/ ils ne parlent pas la langue/ <roma/>
18	M71A : <Ils ne par>lent pas la langue/ langue.
19	Ch : Pourquoi ils ne parlent pas <la langue ?>
20	M71A : <Ils ne par>lent pas la langue.
21	S/ on <les> appelle Roms albanais.
22	Ch : <Pourquoi> ils ne parlent pas <la langue ?>
23	M71A : Ils ne parlent pas parce qu'ils ne sont pas Roms !
24	Ch : Ils ne sont pas <Roms?>
25	M71A : <Non>!
26	Ils ne sont pas non plus Roms.
27	Ch : Hmm.
	[silence]
28	(?) pour eux tous les Égyptiens.
29	Ch : Ils viennent d'Egypt/ <d'Egypte.>
30	M71A : <Egypte/ Eh.>
31	Ch : Quand est-ce qu'ils sont arrivés ? Tu sais toi ?
32	M71A : Eh comme ça, ils sont partis/ aux temps anciens (litt. dans les anciennes années).
33	Les Égyptiens ont fui, le peuple, il y a eu la guerre.
34	Le peuple est parti.
35	(?) ils ont fui comme/ nous Albanais/ quand le peuple a fui/ en Allemagne Italie France/ Angleterre.
36	Ils sont allés partout !
37	Et ils (?) maintenant !
38	On les dit albanais.
39	Eh ah en Grèce il y a/ trois villages où on parle albanais.
40	Ils parlent notre langue/ albanais.

41	Ch : Hmm.
42	M71A : D'où ils viennent ? d'Albanie.
43	Ils sont devenus Grecs.

(173)	*Eh*	*akowa*	*phira-d-e*	*otakal/*	*t-**o***	***puran-e***	***breš-a.***
	Eh	DEM	marcher-PRET-3PL	ainsi	à-ART.DEF.PL	ancien-PL	année-PL

« Eh comme ça, ils sont partis/ il y a longtemps. (litt. dans les anciennes années). »

Statut de l'information

Il s'agit de la première mention de l'époque ancienne dans la conversation : le référent de *to purane breša*, aux temps anciens, est nouveau lexicalement. L'adjectif *anciens* est nouveau lexicalement.

Structure de l'information

Le syntagme de *to purane breša*, dans les anciennes années, constitue le topique-cadre (*frame-setting topic*) de l'expression *phirade ot akal*, ils sont partis.

Interprétation sémantique

Ici 'anciens' fait référence à une époque révolue et située dans une époque lointaine et indéterminée. Le sens global du syntagme est général et non spécifique ; l'adjectif fait partie intégrante de l'expression référentielle globale. *Les anciennes années* fait référence à une époque lointaine car les années qui la séparent du présent sont nombreuses. L'adjectif revêt presque une valeur de quantification, qui vient conférer au référent de *années* la propriété d'être très distante par rapport au moment de l'énonciation. Le syntagme en (173) est donc lui aussi un exemple de modification directe.

(174) M50A, Korçë, 5 août 2014

L'extrait suivant est un entretien dirigé où la le locuteur présente sa famille.

1	*M50A: Oj isine Elbasani me sasa Korča školə.*
2	*Kaž lelom le oj isine/ školasar othe/ nə škol tə mesmes.*
3	*Ch: Tire maksemja đele ti škola?*
4	*M50A: Grećija/ atka na.*
5	*Na džinena/ čhibane ole/ sare grećišt džinena.*
6	*Zavilola ti škola.*
7	*Ch: Aθina.*
8	*M50A: Aθina sar Aθina ti škola isi.*
9	*I tikni čhaj/ ćerdom la othe/ Grećija.*
10	*Ch: Kećі/ breša/ isi i bari isi i tikni?*
11	*M50A: Dešo/ duj brəšengi isi i tikni.*
1	M50A : Elle[1] était à Elbasan et moi à Korçë à l'école.
2	Quand je me suis mariée avec elle (litt. je l'ai prise), elle elle était/ à l'école là/ au lycée.

[1] Son épouse.

3	Ch : Tes enfants ont été à l'école ?
4	M50A : En Grèce/ pas ici.
5	Ils ne parlent pas/ albanais eux/ ils parlent tous grec.
6	(?) à l'école.
7	Ch : A Athènes.
8	M50A : A Athènes biens sûr que c'est à Athènes qu'ils vont à l'école.
9	La petite fille/ je l'ai faite là/ en Grèce.
10	Ch : Combien/ d'années/ a la grande a la petite ?
11	M50A : D/ douze ans elle a la petite.

(174)	***I***	***tik-ni***	***čhaj***	*ćer-d-om*	*la*	*othe/*	*Grećija.*
	ART.DEF.F.SG	petit-F.SG	fille(F.SG)	faire-PRET-1SG	3.SG.F.OBL	là	Grèce

« La petite fille/ je l'ai faite là/ en Grèce. »

Statut de l'information

Le référent « petite fille » n'a encore jamais été mentionné, mais il est accessible-inférable (*accessible-inferable set-relationship*) car il est préparé par « tes enfants » depuis plusieurs tours de parole. L'adjectif *petite* est nouveau lexicalement.

Structure de l'information

Le syntagme *i tikni čhaj*, la petite fille, constitue le topique-propos (*aboutness topic*) de l'expression *ćerdom la othe/ Grećija*, je l'ai faite là/ en Grèce.

Interprétation sémantique

La référence se calcule par l'adjonction des sèmes de l'adjectif et de ceux du nom : le référent de l'expression globale est à l'intersection de l'ensemble {fille} et de l'ensemble {entités petites}. L'identification du référent se fait en opérant une restriction pragmatique au sein des référents du nom. L'adjectif porte un sème spécifique mais pas nécessairement contrastif puisque rien n'est dit du lieu de naissance des autres enfants[1]. On peut cependant considérer le syntagme en (173) comme un exemple de modification indirecte : il diffère des deux exemples précédents.

(175) M46A, Korçë, 22 juillet 2013

Le récit de vie suivant a été raconté au cours d'une conversation spontanée.

1	*M46A: Oke vaćerela/ XXX lesće papus/ lesće mamja/ lesće dades em maksemes džanes.*
2	*Čhudam/ firma/ kaj/ akva čhavo ka ovel operacjoni aman/ sa isi jazmiši/ kato baro Del.*
3	*Ch: ə/*
4	*M46A: Jašar/ na jašar.*
5	

[1] Le locuteur a quatre autres enfants, dont deux sont nés en Albanie d'une première femme, et deux sont nés en Grèce, d'une deuxième femme.

6 7	*Em akale maksemes/ vaćrdas ma o doktori/ le les tu Beni akana te angaljate/ em prastando te inđeles les/ ti/* [le téléphone sonne] *ti vaver spitalja.* *Kaj amen/ kerdam azəri.*
1 2 3 4 5 6 7	M46A : Il dit OK/ XXX son grand-père/ sa grand-mère/ son père et l'enfant tu sais. On a/ signé/ que/ cet enfant soit opéré mais/ tout est écrit/ par Dieu tout puissant. Ch : Euh/ M46A : Ecrit/ pas écrit. Et cet enfant/ me dit le docteur/ prends le dans tes bras maintenant, toi Beni/ et cours l'amener/ à/ [le téléphone sonne] à l'autre hôpital. Parce que nous/ on a fait (?).

(175)	*Sa*	*isi*	*jazmiši/*	*kat-**o***	***bar-o***	***Del.***
	tout	être.3.PRES	écrit	venant_de-ART.DEF.M.SG	grand-M.SG	Dieu(M.SG)

« Tout est écrit/ par Dieu tout puissant. (litt. par Dieu grand) »

Statut de l'information

Il s'agit de la première mention de Dieu dans le récit : le référent de *kato baro Del*, par Dieu tout puissant, est nouveau. Le dieu auquel il est fait référence, et c'est pourquoi nous y avons mis une majuscule, est le dieu chrétien orthodoxe, qui a pour attribut d'être tout puissant. On pourrait objecter que Dieu fait partie de la connaissance du monde attendue, ce qui le rend accessible-général. Dans ce cas, l'adjectif *grand* transmet une propriété elle aussi accessible-général, même si l'adjectif est nouveau lexicalement.

Structure de l'information

Le syntagme *kato baro Del*, par Dieu tout puissant, constitue le topique-cadre (*frame-setting topic*) de *sa isi jazmiši*, tout est écrit – même si Dieu est l'agent de l'action d'écrire.

Interprétation sémantique

L'adjectif *baro*, grand, transmet une propriété du référent qui lui est inhérente, sa puissance. Un dieu est forcément grand. En outre dans une perspective monothéiste, l'adjectif *grand* n'a pas une interprétation intersective avec le nom *Dieu* : il n'y a pas de restriction pragmatique puisqu'il n'y a qu'un seul dieu. L'interprétation est donc hiérarchique, à l'instar du syntagme polydéfini des *cobras venimeux* de Kolliakou (2004). On peut presque parler d'épithète de nature : il s'agit d'un cas où le modifieur, ici l'adjectif, présente très peu de traits sémantiques spécifiques. Le sens de l'adjectif apparaît comme redondant, compris par le sens du nom : la référence de l'expression globale ne se calcule pas par l'adjonction des sèmes de l'adjectif et de ceux du nom. Le syntagme en (175) est donc un exemple de modification directe.

(176) M71A, Korçë, 28 juillet 2014

Le récit de vie suivant a été raconté au cours d'une conversation spontanée.

1	*M71A: Me te/ pišav e gurumna/ te kerava lo thud/ te astarav/ yiurti/ te kerav kiral/ kerava sine.*
2	*XXXX/ kaj isine buljarmiši kaj sovenas.*
3	*Me das/ čhuvas lenge thaneste.*
4	*Ka lavas i me tha te šulava isine/ te keravas mo maro/ te keras amenge kave/ e phurenge.*
5	*Hajde te/ džava sorathan.*
6	*Pišavas e gurumna.*
7	*Kera/ keravava sine o thud/ astaravas jiurti/ jaurti džanesa akava jaurti.*
8	*Čhuvavas odova maja jaurtesele/ kirolas o thud/ mukhas šudrola sine/ em keravas jaurti.*
9	*Sotha maravas les keravas les sine kava sine o kir.*
10	*Jek odole kašteste sine akhal baro.*
11	*Maravas les/ em mukhava sine o kir.*
12	*džanesa o kir so vakeren le? e gurumnako.*
13	*Xa sine šutalo thud/ panesće thaneste pijas sine/ o šutalo thud.*
14	*O jiurti kaj sine/ čhuvavas panik de/ kaj marjola sine ov ovela sine sar/ XXX/ as jaurti as pani.*
15	*Pijavas les e panesće thaneste.*
16	*Isine amen o dobiča!*
17	*Na pijas pani.*
1	M71A : Moi il fallait/ que je traie les vaches/ que je fasse le lait/ que je prépare/ du yaourt/ que je fasse le fromage/ je faisais.
2	XXXX/ qui étaient couchés qui dormaient.
3	Moi je donnais/ je me mettais à leur place.
4	J'y allais moi aussi et je balayais/ pour faire mon pain/ pour nous faire du café/ pour les vieux.
5	Allez pour/ y aller après.
6	Je trayais les vaches.
7	Je fais/ je faisais le lait/ je préparais du yaourt/ yaourt tu connais ça le yaourt.
8	Je mettais ça, de la levure dans le yaourt/ je faisais bouillir le lait/ je le laissais refroidir/ et je faisais du yaourt.
9	Ensuite je le battais je le faisais, ça c'était le beurre.
10	Dans ce bois l'un était grand comme ça.
11	Je le battais/ et je laissais reposer le beurre.
12	Tu sais comment on appelle ça le beurre ? de vache.
13	On mangeait du lait fermenté/ à la place de l'eau on buvait/ du lait fermenté.
14	Le yaourt qu'on avait/ on y mettait de l'eau/ pour qu'il (?), il devenait comme/ XXX/ ni du yaourt ni de l'eau.
15	Je buvais ça à la place de l'eau.
16	On avait les juments !
17	On buvait pas de l'eau.

(176)	*Pan-es-ć-e*	*than-es-te*	*pij-as*	*sine/*	***o***	***šutal-o***	***thud.***
	eau-M.SG.OBL-GEN-M.SG.OBL	lieu-M.SG.OBL-DAT	boire-1PL	être.3.PRET	ART.DEF.M.SG	acide-M.SG	lait(M.SG)

« Au lieu de l'eau on buvait/ du lait fermenté. »

Statut de l'information

Le référent de *o šutalo thud*, le lait fermenté, est donné-actif puisqu'il vient d'être mentionné dans la même phrase. Le nom *lait* comme l'adjectif *acide* sont donnés lexicalement.

Structure de l'information

Le syntagme *o šutalo thud*, le lait fermenté, constitue, avec le verbe *pijas sine*, on buvait, le focus contrastif (*contrastive subtype replacing*, selon la typologie de Götze et al. 2007) de *panesće thaneste*, à la place de l'eau.

Interprétation sémantique

L'adjectif *šutalo*, acide, transmet une propriété du référent qui lui est inhérente, son acidité. Ici le référent n'est pas à l'intersection de l'ensemble {lait} et de l'ensemble {entités acides}, car il n'est pas question de lait non-acide. La référence ne se calcule par la simple adjonction des sèmes de l'adjectif 'acide' et de ceux du nom 'lait' : il s'agit d'un type de boisson qui ne s'oppose pas à du lait non-acide, mais, en bloc, à de l'eau. Le référent est proche du kéfir ou du lait ribot. La séquence A+N fonctionne comme un nom composé. En outre, le référent lait fermenté a un sens générique dans cet extrait. Le syntagme en (175) est donc un exemple de modification directe.

(177) – (178) F44AČ, F65A, Korçë, 1er août 2014

L'échange suivant est issu d'une conversation familiale spontanée, consacrée à différents types d'esprits.

1	*Ch: ə/ kon šaj te ovel ə/ vampiri?*
2	*Ko/ sar manuša?*
3	*F65A: O bilače.*
4	*Sa o bilače.*
5	*F44AČ: O bilačho manuš ovela vampiri.*
6	*F65A: O bilačho/ ah/*
7	*F85A: <Akova but bilačho vampiri te ovel.>*
8	*F65A: <O bilačho ođesko.*
9	*Ov> kerdas but bilačhipe/ but manušen mudardas/ ə/ but ə/ čordas.*
10	*But bilačhipe kerdas.*
11	*<Akova.>*
12	*Ch: <O lače ma>nuša naštin te oven/ <vampiri.> <Vampiri?>*
13	*F65A: <Nna.> O lače manu<ša> džala lengo ogi to devel!*
1	Ch : Euh/ qui peut devenir euh/ vampire ?
2	Qu/ quels gens ?
3	F65A : Les méchants.
4	Tous les méchants.
5	F44AČ : Le méchant homme devient vampire.
6	F65A : Le méchant/ ah/
7	F85A : <Celui-ci deviendra un vampire très méchant.>
8	F65A : <Celui qui a méchant cœur.
9	Lui> il a fait beaucoup de mal/ il a tué beaucoup de gens/ euh/ il a beaucoup euh/ volé.
10	Il a fait beaucoup de mal.
11	<Lui.>
12	Ch : <Les gens (litt. hommes) gentils> ils ne peuvent pas devenir/ <vampires.> <Vampires?>
13	F65A : <Nnon.> Les <gens> (litt. hommes) gentils, leur âme va au ciel !

(177)	***O***	***bilačh-o***	***manuš***	*ov-ela*	*vampiri.*
	ART.DEF.M.SG	méchant-M.SG	homme(M.SG)	devenir-3SG.PRES	vampire

« Le méchant homme devient vampire. »

(178)	***O***	***lačh-e***	***manuša***	*dž-ala*	*leng-o*	*ogi*	*to*	*devel!*
	ART.DEF.PL	bon-PL	homme(M.PL)	aller-3SG.PRES	POSS.3PL-M.SG	esprit(M.SG)	à.ART.DEF.M.SG	ciel(M.SG)

« Les gens (litt. hommes) gentils, leur âme va au ciel ! »

Statut de l'information

Le référent « homme méchant » est donné-actif référentiellement, puisqu'il vient d'en être question dans le tour de parole précédent. L'adjectif *méchant* comme le nom *homme* sont donnés lexicalement.

Le référent « hommes gentils » est nouveau. L'adjectif *gentils* est nouveau lexicalement – mais on pourrait arguer qu'il est accessible-inférable puisqu'il entre dans une opposition binaire avec *méchant*, qui a été prononcé plusieurs fois. Le nom *hommes* est donné-inactif lexicalement car il a été mentionné plus haut dans le discours.

Structure de l'information

Le syntagme *o bilačho manuš*, le méchant homme, constitue le topique-propos (*aboutness topic*) de l'expression *ovela vampiri*, devient vampire.

Le syntagme *o lače manuša*, les hommes gentils, constitue, avec *lengo ogi*, leur âme, le topique-propos (*aboutness topic*) de l'expression *džala to devel*, va au ciel.

Interprétation sémantique

La référence se calcule par l'adjonction des sèmes de l'adjectif et de ceux du nom : les référents des expressions globales sont à l'intersection de l'ensemble {hommes} et de l'ensemble {entités méchantes} ou {entités gentilles}. Les deux référents font l'objet d'une opposition binaire. Mais les deux adjectifs sont en opposition frontale de manière générique et non spécifique. Le profil général du *bon* s'oppose au profil du *méchant*. L'adjectif est mis au premier plan *avec* le nom, ce qui produit un effet de mariage sémantique (Nølke 2001 : 181). Il y a donc contraste, quand bien même il y a modification directe et il n'y a pas focus étroit sur l'adjectif.

(179) – (201) M46A, Korçë, 22 juillet 2013

Ce récit de vie a été raconté au cours d'une conversation spontanée.

1	*M46A: Akana/ lesće mami akale/ čhavesći em o papu leskoro/ xana pes e doktorensar/ kaj te kurtaren le/ lenge maksemes.*
2	*ə/ manđe bičhalena Beni to štar te ranoste.*

3	*Kan džava dikhav akale maksemes/ em ćerava lafi e bare doktorensar.*
4	*O baro/ doktori vaćerela mange: "Beni/ akale maksemesće amen/ so isine ćerimeasće ćerdam les.*
5	*Akale maksemese te čhuvel o Del vast/ te kurtarel les.*
1	M46A : Maintenant/ sa grand-mère de ce/ garçon et son grand-père/ ils se battaient avec les docteurs/ pour le guérir le/ leur enfant.
2	Euh/ à moi Ben ils m'envoient à quatre heures du matin.
3	Quand j'arrive je vois cet enfant/ et je discute avec les grands docteurs.
4	Le grand/ docteur me dit : « Beni/ pour cet enfant nous/ tout ce qu'il y avait à faire on l'a fait.
5	Cet enfant il faut que Dieu l'aide/ à guérir.

(179)	*em*	*ćer-ava*	*lafi*	***e***	***bar-e***	***doktor-en-sar***.
	et	faire-1SG.PRES	parole	ART.DEF.PL.OBL	grand-PL.OBL	docteur-M.PL.OBL-INSTR

« […] et je discute avec les grands docteurs. »

Statut de l'information

Le référent de « avec les grands/ docteurs » est donné-inactif car le locuteur les a mentionnés deux phrases plus haut. L'adjectif est lexicalement nouveau.

Structure de l'information

Le syntagme *e bare doktorensar*, avec les grands/ docteurs, constitue un focus de nouvelle information non sollicitée, avec le verbe *ćerava lafi*, je discute.

Interprétation sémantique

L'adjectif *bare*, grands, transmet une propriété du référent « docteurs » qui lui est quasiment inhérente, sa puissance. Le locuteur veut signifier que l'enfant a été soigné par des personnes compétentes, des spécialistes. C'est pourquoi l'adjectif a selon nous une valeur emphatique et non contrastive : il n'est nulle part question de « petits docteurs » ou de « mauvais docteurs ». Ceci va de pair avec le fait que l'enfant soit soigné à l'hôpital et non chez lui, et qu'il s'agisse de l'hôpital de la capitale, Tiranë, et non de celui de sa région, Korçë. On peut presque parler d'épithète de nature : il s'agit d'un cas où le modifieur, ici l'adjectif, présente très peu de traits sémantiques spécifiques. Le syntagme est un exemple de modification directe.

(201)	***O***	***bar-o/***	***doktori***	*vaćer-ela*	*man-ge*	[…]
	ART.DEF.M.SG	grand-M.SG	docteur(M.SG)	parler-3SG.PRES	1SG.OBL-DAT	

« Le grand/ docteur me dit […]. »

Le syntagme (201) présente une pause de 0,44 seconde.

Statut de l'information

Le référent de « le grand/ docteur » est donné-actif car il vient d'être mentionné à la phrase précédente. L'adjectif est donné-actif lexicalement.

Structure de l'information

Le syntagme *o baro/ doktori*, le grand docteur, constitue le topique-propos de *vaćerela mange*, me dit.

Interprétation sémantique

Comme pour l'exemple (179), on peut presque parler d'épithète de nature : il s'agit d'un cas où le modifieur, ici l'adjectif, présente très peu de traits sémantiques spécifiques. Le syntagme est un exemple de modification directe.

2.7. Corpus spontané : DEF N A

(180) F16AČ, Korçë, 22 juillet 2014

L'exemple suivant est issu d'une recette de cuisine racontée par la fille de mon hôte.

1	*F16AČ: Ah/ jek vaver isi/ kaj ćerasa o/ phiravasa o pete bare.*
2	*Phiravasa len.*
3	*Ch: So isi o pete?*
4	*F16AČ: O pete isi sar ə/ sar ćerdan tu/ Auror/ e/ me dajasar marikli/ plečinta/ kaj ćerdan*
5	*marikli.*
6	*Ch: O plečinta?*
7	*F16AČ: Va.*
8	*Na odole! Kaj phiradan.*
9	*ə/ oklajasar kaj phiradan.*
10	*ə/ lasa odole pete em čhuwasa len/ jek čarčafeste avrik.*
11	*Phiravasa len to/ to/ divanja.*
12	*Phiravasa len.*
13	*Em ə/ šućona.*
14	*Em/ šućona em ovena tikne.*
1	F16AČ : Ah/ une autre c'est/ quand on fait les/ on pétrit les grosses boules de pâte.
2	On les pétrit.
3	Ch : C'est quoi les *boules de pâte* ?
4	F16AČ : Les boules de pâte c'est comme euh/ comme tu as fait toi/ Aurore/ avec/ avec ma
5	mère la marikli/ la plečinta/ quand tu as fait la marikli.
6	Ch : Les plečinta?
7	F16AČ : Oui.
8	Pas celles-là! Que tu as pétri.
9	Euh/ avec celui-là que tu as pétri.

10	Euh/ on prend ces boules de pâte et on les met/ dans un torchon dehors.
11	On les pétrit sur le/ sur le/ canapé.
12	On les pétrit.
13	Et euh/ elles sèchent.
14	Et/ elles sèchent et elles deviennent petites.

(180)	*Phirav-asa*	***o***	***pet-e***	***bar-e***.
	pétrir-1PL.PRES	ART.DEF.PL	boule_de_pâte-F.PL	grand-PL

« On pétrit les grosses boules de pâte. »

Statut de l'information

Il s'agit du début de la recette, donc de la première mention du référent « grosses boules de pâte » : il est nouveau. La locutrice vient me donner d'autres recettes de cuisine qui mettaient en jeu de la pâte, mais elle n'a jusqu'alors jamais employé le terme *pete*, qui est nouveau lexicalement. L'adjectif *grands* est lui aussi nouveau lexicalement.

Structure de l'information

Le syntagme *o pete bare*, les grosses boules de pâte, constitue avec le verbe le focus de la phrase, plus précisément un focus de nouvelle information non sollicitée (*unsolicited new-information focus*).

Interprétation sémantique

Pour réaliser la recette, il importe de prendre de grosses boules de pâte, car lorsqu'on les met à sécher, elles se réduisent – information que l'on n'apprend qu'à la fin de la recette. Afin que les boules soient petites à la fin du processus, il importe qu'elles soient initialement grandes, d'où une attention portée à leur dimension. Cette attention se traduit syntaxiquement par la postposition de l'adjectif. Un contraste implicite existe, à l'échelle de toute la recette, entre *bare*, grandes, et *tikne*, petites : le même référent évolue au fil de la recette. Ce contraste étant implicite, il est réalisé *in absentia*. Notons que s'il existe un focus étroit à valeur contrastive sur l'adjectif, seule la locutrice en est consciente mais non l'interlocutrice qui ignore la suite des opérations.

(181) – (182) – (196) M58MČ, Voskopojë, 15 juillet 2013

Dans l'entretien dirigé suivant, le locuteur présente ses enfants.

1	*Ch: Tiri familja?*
2	*M58MČ: A, mərəmika. ə/ hmm.*
3	*Man si man duj čhave/ jek che.*
4	*So trin si ʒenime/ ə/*
5	*Ch: Bare si?*
6	*M58MČ: Si bare.*

7	*ə jek si/ ə trijanda enja bəršengo.*
8	*Jek si trijanda/ em pandž bəršengo.*
9	*E čhej kaj si ma i bari/ si saranda jek bəršeng.*
10	*Ch: Hmm.*
11	*M58MČ: ə/ kə čhej si la/ duj cinonə.*
12	*ə/ e čhave bare si le/ ə duj čheja jek čhavo.*
13	*ə o čhavo cikno/ murno čhavo kaj si/ ə/ trijanda pandžə bəršengo/ ə/ vo ə/ lje si le jek čhe.*
1	Ch : Ta famille ?
2	M58MČ : Ah, (?). Euh/ hmm.
3	J'ai deux fils/ une fille.
4	Tous les trois sont mariés/ euh/
5	Ch : Ils sont grands ?
6	M58MČ : Ils sont grands.
7	Euh un a/ euh trente et un ans.
8	Un a trente/ cinq ans.
9	La fille que j'ai la grande/ a quarante et un ans.
10	Ch : Hmm.
11	M58MČ : Euh/ la fille elle a/ deux petits.
12	Euh/ le grand garçon il a/ euh deux filles un fils.
13	Euh le petit garçon/ mon fils qui a/ euh/ trente cinq ans/ euh/ lui euh/ lui il a une fille.

(181)	*ə/*	***e/***	***čhav-e***	***bar-e***	*si*	*le/*
	euh	ART.DEF.M.SG.OBL	fils-M.SG.OBL	grand-M.SG.OBL	être.3.PRES	3PL
	ə	*duj*	*čhej-a*	*jek*	*čhavo.*	
	euh	deux	fille-F.PL	un	fils(M.SG)	

« Euh/ le grand garçon il a/ euh deux filles un fils. »

(182)	*ə*	***o***	***čhavo***	***cikn-o/***	*murn-o*	*čhavo*	[...]
	euh	3SG.M	fils(M.SG)	petit-M.SG	POSS1SG-M.SG	fils(M.SG)	
	vo	*ə/*	*lje*	*si*	*le*	*jek*	*čhe.*
	3SG.M	euh	3SG.M.OBL	être.3.PRES	3SG.M.OBL	un	fille(F.SG)

« Euh le petit garçon/ mon fils qui a/ euh/ trente cinq ans/ euh/ lui euh/ lui il a une fille. »

Statut de l'information

Le référent de *e čhave bare*, le grand garçon, est accessible-inférable (set relation) puisqu'il a annoncé un peu plus haut avoir deux fils. L'adjectif *bare* est donné-inactif lexicalement car il a été mentionné plus haut dans la conversation.

Le référent de *o čhavo cikno*, le petit graçon, est accessible-inférable (set relation). Le nom *čhavo*, garçon, est donné-actif puisqu'il vient d'être mentionné dans la phrase précédente pour deux autres référents. L'adjectif *cikno*, petit, est donné-inactif lexicalement, car il a été mentionné deux phrases plus haut pour d'autres référents.

Structure de l'information

Le syntagme *e čhave bare*, le grand garçon, constitue le topique-propos (*aboutness topic*) de l'expression *si le/ ə duj čheja jek čhavo*, il a/ euh deux filles un fils.

Le syntagme *o čhavo cikno*, le petit garçon, constitue le topique-propos (*aboutness topic*) de l'expression *lje si le jek čhe*, lui il a une fille. Les deux phrases sont symétriques.

Interprétation sémantique

La référence des deux syntagmes se calcule chaque fois par l'adjonction des sèmes de l'adjectif et de ceux du nom : le référent de l'expression globale est à l'intersection de l'ensemble {fils} et de l'ensemble {entités grandes} et {entités petites}. L'identification du référent se fait en opérant une restriction pragmatique au sein des référents du nom. Les deux adjectifs portent un sème spécifique : le référent de *e čhave bare*, le grand garçon, est contrastif du référent de *o čhavo cikno*, le petit garçon, mentionné dans la phrase suivante. Le contraste se fait *in presentia*. Il existe donc un focus étroit à valeur contrastive portant sur les adjectifs *bare*, grands, et *cikno*, petit. Les syntagmes (181) – (182) sont des exemples de modification indirecte.

(196)	***E***	***čhej***	*kaj*	*si*	*ma*	***i***	***baɻ-i/***
	ART.DEF.F.SG	fille(F.SG)	REL	être.3.PRES	1SG.OBL	ART.DEF.F.SG	grand-F.SG
	si	*saranda*	*jek*	*bərš-en-g.*			
	être.3.PRES	quarante	un	an-M.PL.OBL-GEN			
	« La fille que j'ai la grande/ a quarante-et-un ans. »						

Le syntagme présente une insertion mais pas de pause.

Statut de l'information

Le référent est donné-inactif car le locuteur a déjà mentionné ses enfants. Le nom est donné-inactif, mais l'adjectif est nouveau lexicalement.

Structure de l'information

Le syntagme *e čhej kaj si ma i bari*, la grande fille/ que j'ai/, constitue le topique-propos de *si saranda jek bəršeng*, a quarante et un ans.

Interprétation sémantique

La partie nominale, la fille, fait partie du présupposé. Le syntagme vise à sélectionner un référent au sein de l'ensemble {enfants}, grâce à l'adjectif *grande*. En effet le référent ne se situe pas au sein de l'ensemble {filles} car celui-ci ne contient qu'une seule entité, le référent. Il s'agit d'un syntagme polydéfini à effet de sélection, où l'adjectif est un modifieur direct. Comme dans l'exemple précédent, la rupture se manifeste par un silence ainsi que par une insertion, ici d'une proposition relative. Elle

est un modifieur supplémentaire qui aide à identifier le référent, à l'instar du nom propre apposé en (195). Le contraste se fait *in presentia* car les différents enfants ont été mentionnés en discours.

(183) F16AČ, Korçë, 24 juillet 2014

L'extrait suivant est issu d'une conversation spontanée.

1	*F16AČ: Asjek khorak na dikhlom diskriminacja.*
2	*Ch: He/ si but šukar.*
3	*Te ovel ak/ akhal andi sari Europa.*
4	*F16AČ: Eh/ isi isi odole/ Vlaxura/ isi Vlaxura kaj nanaj diskriminacja/ sar isi akale komšiji akate/ akale vavera.*
5	*Hm/ sosće akale ba/ barisajle to/ sar to Roma/ ə/ to amalinja vavera/ akale bar/ barisajle to breša/ to veš.*
6	*Em/ me phenlom lenge/ tumen isi/ isi/ isi tumen but bari diskriminacja sosće/ ni džanen/ so isi diskriminacja.*
1	F16AČ : Pas une seule fois j'ai vu de la discrimination.
2	Ch : Eh/ c'est super.
3	Je souhaite que ce soit co/ comme ça dans toute l'Europe.
4	F16AČ : Eh/ il y a il y a ces/ Vlax/ il y a les Vlax qui n'ont pas de discrimination/ comme sont ces voisins ici/ ces autres.
5	Hmm/ parce qu'ils ont/ ils ont grandi dans/ comme chez les Roms/ euh/ chez les autres amies/ ceux-ci ils ont/ ils ont grandi avec les années/ dans la forêt.
6	Et/ moi je leur ai dit/ vous vous avez/ avez/ vous subissez une grande discrimination parce que/ vous ne savez pas/ ce que c'est que la discrimination.

(183)	*Sosće*	*akale*	*ba/*	*bar-isajl-e*	*to/*	*sar*
	parce_que	DEM.PL	gran/	grandir-PRET-3PL	à-ART.DEF.PL	comme
	to	*Rom-a/*	*t-**o***	***amal-in-ja***	***vaver-a.***	
	à-ART.DEF.PL	Rom-PL	à-ART.DEF.PL	ami-F-PL	autre-PL	

« Parce que ceux-ci ils ont/ ils ont grandi dans/ comme chez les Roms/ euh/ chez les autres amies. »

Statut de l'information

Le référent « les autres amies » n'a encore jamais été mentionné dans cette conversation, il est nouveau. L'adjectif *autres* est nouveau lexicalement.

Structure de l'information

Le syntagme *to amalinja vavera*, chez les autres amies, constitue, avec *barisajle to/ sar to Roma/*, ils ont grandi dans/ comme chez les Roms/, le focus de nouvelle information.

Interprétation sémantique

La référence se calcule par l'adjonction des sèmes de l'adjectif et de ceux du nom : le référent de l'expression globale est à l'intersection de l'ensemble {amies} et de l'ensemble {autres entités}. L'identification du référent se fait en opérant une restriction pragmatique au sein des référents du nom. Le syntagme en (183) est un exemple de modification indirecte. Aucune amie n'est présente ni a été mentionnée précédemment : le contraste se fait *in absentia*.

(184) M71A, Korçë, 27 juillet 2013

L'extrait suivant est issu d'une conversation spontanée.

1	*M71A: Oh bisterdam amen/ bisterdam but breša akale/ gilja/ e ə/ purane gilja but.*
2	*Ko šele brešengi gilja si akale.*
3	*Giljavena sa angle o phure akale gili.*
4	*Koraxane gilja majna sine/ o phure/ but koraxane gilja nas*ʔ*.*
5	*Ta vaćeras romane/ elenika giljave<na sine.>*
6	*Ch: <Ah.>*
7	*M71A: O bo but koraxane giljavena s amare phure.*
8	*Ch: Hmm.*
9	*Džanesa tu?*
10	*M71A: Ka đilava len me! Akana bisterdom odole gilja koraxane.*
1	M71A : Oh on a oublié nous/ on a oublié depuis de nombreuses années ces/ chansons/ eh euh/ chansons très anciennes.
2	Elles ont à peu près cent ans ces chansons.
3	Ce sont les vieux qui les chantaient jadis, ces chansons.
4	C'était des chansons turques tu vois/ les vieux/ il y avait (?) beaucoup de chansons turques.
5	Et on parlait en romani/ ils chant<aient> en grec.
6	Ch : <Ah.>
7	M71A : Bah oui nos vieux ils chantaient beaucoup de chansons turques.
8	Ch : Hmm.
9	Tu en connais toi ?
10	M71A : Que je les chante moi ! Maintenant j'ai oublié ces chansons turques.

(184)	*Bister-d-om*	***odole***	***gil-ja***	***koraxan-e***.
	oublier-PRET-1SG	DEM.F.SG	chanson-F.PL	musulman-PL

« Maintenant j'ai oublié ces chansons turques. »

Statut de l'information

Le référent « chansons turques » a déjà été mentionné dans cette conversation : le syntagme nominal est donné-inactif référentiellement, de même le nom et l'adjectif lexicalement. Si l'on ignore le tour de parole de l'interlocutrice, tout est donné-actif du point de vue du locuteur.

Structure de l'information

Le syntagme *odole gilja koraxane*, ces chansons turques, constitue avec le verbe le focus de la phrase, plus précisément un focus de nouvelle information sollicitée (*solicited new-information focus*).

Interprétation sémantique

Le locuteur a commencé par mentionner les chansons turques que chantaient les anciens[1], avant de mentionner le répertoire romani chanté en grec, et de conclure que les chansons turques étaient plus fréquemment chantées... S'il y a contraste avec des chansons en grec, il est réalisé *in presentia*. Si le contraste se fait avec des chansons (plus récentes) en albanais, alors il est réalisé *in absentia*. Lorsqu'il répond les avoir oubliées, il emploie un démonstratif déictique du discours : il fait référence à ces chansons musulmanes précises, celles connues et préférées par les anciens en question. Comme dans l'exemple (180), la référence se calcule par l'adjonction des sèmes de l'adjectif et de ceux du nom : le référent de l'expression globale est à l'intersection de l'ensemble {chansons} et de l'ensemble {entités turques}. L'identification du référent se fait en opérant une restriction pragmatique au sein des référents du nom. Notons que l'adjectif porte d'autres sèmes outre celui de la nationalité : un sème relatif aux ancêtres, un sème relatif aux préférences de ceux-ci. Il est donc très spécifique, ce qui permet le rend très aisément identifiable. Le syntagme en (184) est lui aussi un exemple de modification indirecte.

(185) M58MČ, Voskopojë, 14 juillet 2013

L'extrait suivant est issu d'une conversation spontanée.

1	*F24AČ: Šiiil?* [rire]
2	*Ch: Na si man šil.*
3	*F24AČ: Naj tut?*
4	*Ch: Na.*
5	*M58MČ & F24AČ:* [discussion en albanais]
6	*F24AČ: Obobobobo. Studenta studenta.*
7	*M58MČ: Ano foro baro Tirana/ but tatipe.*
8	*Ch: Ah.*
9	*M58MČ: But.*
10	*Ch: Akate si mišto.*

[1] Notons un problème d'accord, dans le syntagme *ko šele brešengi gilja*, entre le nom (qui est au pluriel) et le modifieur, un groupe génitival (qui est au singulier). On attendrait en effet *ko šele brešenge gilja*. Nous rencontrons un problème de morphologie similaire dans l'exemple (261), cf. note 1, p. 223.

11	*M58MČ: Ah mišto/ XXX.*
1	F24AČ : T'as froid ? [rire]
2	Ch : Je n'ai pas froid.
3	F24AČ : T'as pas froid ?
4	Ch : Non.
5	M58MČ & F24AČ : [discussion en albanais]
6	F24AČ : Obobobobo. Étudiante étudiante.
7	M58MČ : Dans la grande ville à Tiranë/ il fait très chaud.
8	Ch : Ah.
9	M58MČ : Très.
10	Ch : Ici on est bien.
11	M58MČ : Ah bien/ XXX.

(185)	*An-**o***	***foro***	***bar-o***	*Tirana/*	*but*	*tatipe.*
	dans-ART.DEF.M.SG	ville(M.SG)	grand-M.SG	Tiranë	beaucoup	chaleur(M.SG)

« Dans la grande ville à Tiranë/ il fait très chaud. »

Statut de l'information

Le référent « grande ville » n'a encore jamais été mentionné dans cette conversation – pas même dans l'échange en albanais : le référent est nouveau. L'adjectif *grande* est nouveau lexicalement. On pourrait objecter que Tiranë fait partie de la culture générale, elle serait alors en réalité accessible-général si l'on suit la typologie établie par Götze et al. (2007).

Structure de l'information

Le syntagme *ano foro baro*, dans la grande ville, constitue le topique-cadre (frame-setting topic) de l'expression *but tatipe*, il fait très chaud, qui est une prédication nominale.

Interprétation sémantique

La référence se calcule par l'adjonction des sèmes de l'adjectif et de ceux du nom : le référent de l'expression globale est à l'intersection de l'ensemble {ville} et de l'ensemble {entités grandes}. L'identification du référent se fait en opérant une restriction pragmatique au sein des référents du nom. L'adjectif apporte ici un sème spécifique qui sert à identifier le référent. La juxtaposition du nom propre Tiranë rend *a posteriori* la référence encore plus spécifique et aisément identifiable. Le syntagme en (185) est lui aussi un exemple de modification indirecte.

Le locuteur réside justement dans la capitale. Mais cette conversation a lieu en plein été, le 14 juillet 2013, dans un café situé dans un bois dans les hauteurs de Moscopole/Voskopojë, ville de petite montagne. Un contraste existe, dans le contexte d'énonciation, entre *la grande ville* et le lieu où se tient la conversation. Ce contraste se fait *in presentia*. Impossible de savoir en revanche si le locuteur entend opposer la ville à la campagne (le café dans les bois) ou bien la métropole à la petite ville provinciale (Moscopole/Voskopojë)… donc de savoir si le focus à valeur contrastive porte sur l'ensemble N+A ou seulement sur A.

2.8. Corpus spontané défini et indéfini : autres exemples

(187) M46A, Korçë, 8 août 2014

L'extrait suivant est issu d'un entretien dirigé.

1	*M46A: Te nikav jek ə/ jek ə/ ə jek ruba/ karakteristiko romandikano/ to vaćeribe řomano na čhibani čhib.*
2	*Te čhuvel sa sa řomani čhib.*
3	*Me ka bešav ka sikaav len/ sar te mothon/ taam romani čhib.*
1	M46A : Pour que je produise un euh/ un euh/ euh un (?)/ caractéristique romani/ dans le parler romani pas en langue albanaise.
2	Pour qu'il mette tout tout en langue romani.
3	Moi je vais prendre le temps qu'il faut je vais leur montrer/ comment parler/ bien(?) en langue romani.

(187)	*Te*	*nik-av*	*jek*	*ə/*	*jek*	*ə/*
	COMPL	sortir-1SG.PRES	IND	euh	IND	euh
	ə	***jek***	***ruba/***	***karakteristik-o***	***romandikan-o/***	***t-o***
	euh	IND	?(F.SG)	caractéristique-M.SG	romani-M.SG	à.ART.DEF.M.SG
	vaćeribe	***řoman-o***	*na*	*čhiban-i*	*čhib.*	
	parler(M.SG)	romani-M.SG	NEG	albanais-F.SG	langue(F.SG)	

« Pour que je produise un euh/ un euh/ euh un (?)[1]/ caractéristique romani/ dans le parler romani pas en langue albanaise. »

Statut de l'information

Le référent « le parler romani » est nouveau dans la conversation. L'adjectif *řomano* est cependant accessible-inférable lexicalement, puisque son synonyme *romandikano* est prononcé juste avant lui.

Le référent « la langue albanaise » est nouveau dans la conversation. Le nom *čhib*, langue, est accessible-inférable puisqu'il a déjà été question de parler, mais il est nouveau lexicalement. L'adjectif *čhibano*, albanais est nouveau lexicalement.

[1] Je suppose qu'en tant que journaliste, il parle d'un film ou d'un documentaire.

Structure de l'information

Le syntagme *to vaćeribe řomano*, dans le parler romani, constitue avec *na čhibani čhib*, pas en langue albanaise, le focus de la phrase. Ils sont tous deux en focus contrastif du type *subtype selection* chez Götze et al. (2007).

Interprétation sémantique

La référence se calcule par l'adjonction des sèmes de l'adjectif et de ceux du nom : les référents des expressions globales sont à l'intersection de l'ensemble {parler} ou {langue} et de l'ensemble {entités romani} ou {entités albanaises}. Les deux référents « le parler romani » et « la langue albanaise » font l'objet d'une opposition binaire. Un focus étroit est porté sur l'adjectif dans les deux cas, cependant le locuteur choisit de mettre en avant *romani* plutôt qu'*albanaise*, qui fait l'objet d'une négation. Le locuteur postpose l'adjectif afin de mettre en avant la propriété de l'adjectif *romano* : le fait que le documentaire soit en langue romani pourrait étonner certains, d'où sa mise en valeur syntaxique et sémantique. Au contraire, l'antéposition de l'adjectif semble refléter cette mise au second plan de l'albanais, comme si la syntaxe venait ici appuyer le sens de la phrase. L'adjectif constitue un exemple de modification indirecte. Un contraste est opéré *in presentia*.

(188) M68A, Voskopojë, 15 juillet 2013

Le récit suivant, à propos de l'Union R·romani Internationale, est issu d'une conversation spontanée qui suit un entretien dirigé.

1	*M68A: Ajal avela/ amari isorija kaj sikadam la/ kato kongreso/ eftavardešujek/ kongreso o jekto e Romengo kaj ondilo/ ti/*
	[pause]
2	*ə/ o jekto kongreso ondilo/ Londra/ to/ to privi eftavardešujek.*
3	*ə/ me/ areslom/ kaj/ đerdom/ to/ maškare/ amare/ ćidimjase ko Roma/ ko kongresi/ ko pandž/ kaj ondilo/ Čekija.*
4	*Ch: Praha.*
5	*M68A: Čekja. Praga.*
6	*O kongreso i/ šov/ ondilo/ ande/ Italja.*
7	*O kongreso efta/ e Romengo I/ IRU internacjonal ondilo ə/ eftato kongreso ondilo/ Kroacija.*
8	*O/ me sinom akana ti organizata/ Amari Phen/ kaj ondilo o kongresi jekto kongresi e*
9	*Romengo/ kaj/ kaj ondilo/ Beograd/ kaj si/ o kongreso amare organizata Amari Phen.*
1	M68A : C'est ainsi que fut/ notre histoire telle qu'on l'apprend/ du congrès/ en soixante-et-onze/ le premier congrès des Roms qui a eu lieu/ à/
	[pause]
2	Euh/ le premier congrès a eu lieu/ à Londres/ au/ au (?) soixante-et-onze.
3	Euh/ moi/ je suis arrivée/ que/ j'étais allé/ aux/ parmi/ nos/ réunions aux Roms/ au congrès/ au cinquième/ qui a eu lieu/ en Tchéquie.
	Ch : A Prague.
4	M68A : En Tchéquie. À Prague.
5	Le congrès ès/ six/ a eu lieu/ en/ Italie.

6	Le congrès sept/ des Roms de l'I/ IRU internationale a eu lieu euh/ le septième congrès a eu
7	lieu/ en Croatie.
	Le/ moi je suis maintenant dans l'organisation/ Notre Sœur/ qui fait le congrès premier
8	congrès des Roms/ qu/ qui a eu lieu/ à Belgrade/ qui est/ le congrès de notre organisation
9	Notre Sœur.

(188)	*ə/*	***o***	***jekto***	***kongreso***	*ondil-o/*	*Londra*
	euh	ART.DEF.M.SG	premier	congrès(M.SG)	être.PRET-3.SG	Londres

« Euh/ le premier congrès a eu lieu/ à Londres […]. »

Statut de l'information

Le référent « le première congrès » est donné-actif dans la conversation. Le nom est donné lexicalement, mais l'adjectif *řomano* est nouveau.

Structure de l'information

Le syntagme *o jekto kongreso*, le premier congrès, constitue le topique-propos de la phrase.

Interprétation sémantique

L'adjectif *jekto*, premier, est un numéral ordinal. La référence de l'expression globale ne se calcule pas par l'adjonction des sèmes de l'adjectif et de ceux du nom. L'adjectif est un exemple de modification directe. Cela n'empêche pas un contraste implicite avec les autres numéraux prononcés par la suite : *pandž*, cinq, *šov*, six, *efta*, sept et *eftato*, septième.

(189) – (190) M27A, Korçë, 5 août 2014

L'échange suivant est issu d'un entretien dirigé où le locuteur présente sa famille.

1	*Ch: Isi tut phrala phenja?*
2	*M27A: Isi man duj phrala/ me dadestar/ na me datar se mo dat prandisala duj fora.*
3	*Si/ po bare mandar o baro phral isi sarandupandž e brešengo.*
4	*O/ vaver phral isi saranda brəšengo.*
5	*Isi po bare mandar.*
6	*Isi man e/ trin phenja sora kaj isinam/ jek ratestar/ prandime.*
7	*Jek isi Pogradeci duj si akətka/ Korča/ prandime.*
8	*Ch: Isi len maksemja?*
9	*M27A: Isi len.*
10	*Jek bešela Italija/ jetonela Italija jek/ jetəl/ nela/ Grećija/ jek akətka Korča.*
1	Ch : Tu as des frères et sœurs ?
2	M27A : J'ai deux frères/ du côté de mon père, pas du côté de ma mère car mon père s'est marié deux fois.
3	Ils sont/ plus grands que moi le grand frère a quarante-cinq ans.

4	Le/ l'autre frère a quarante ans.
5	Ils sont plus grands que moi.
6	J'ai aussi/ trois sœurs ensuite qu'on est/ du même sang/ mariées.
7	Une est à Pogradec deux sont ici/ à Korçë/ mariées.
8	Ch : Elles ont des enfants ?
9	M27A : Oui.
10	Une habite en Italie/ elle vit en Italie une/ v/ it/ en Grèce/ une ici à Korçë.

(189)	*Si/*	*po*	*bar-e*	*man-dar*	***o***	***bar-o***
	être.3.PRES	COMP	grand-PL	SG.OBL-ABL	ART.DEF.M.SG	grand-M.SG
	phral	*isi*	*sarandupandž*	*e*	*breš-en-g-o*	
	frère(M.SG)	être.3.PRES	quarante-cinq	ART.DEF.PL.OBL	année-M.PL.OBL-GEN-M.SG	

« Ils sont/ plus grands que moi le grand frère a quarante-cinq ans. »

(190)	***O/***	***vaver***	***phral***	*isi*	*saranda*	*brəš-en-g-o.*
	ART.DEF.M.SG	autre[M.SG]	frère(M.SG)	être.3.PRES	quarante	année-M.PL.OBL-GEN-M.SG

« Le/ l'autre frère a quarante ans. »

Statut de l'information

Le référent de *o baro phral*, le grand frère, est donné actif : il est inclus dans *duj phrala*, deux frères, à la phrase précédente. L'adjectif *baro* est donné lexicalement puisque l'information vient d'être donnée au sein de la même phrase.

Le référent de *o vaver phral*, l'autre frère, est donné inactif : il est inclus dans *duj phrala*, deux frères, deux phrases plus haut. L'adjectif *vaver* est nouveau lexicalement.

Structure de l'information

Le syntagme *o baro phral*, le grand frère, constitue le topique-propos (*aboutness topic*) de l'expression *isi sarandupandž e brešengo*, a quarante-cinq ans.

Le syntagme *o vaver phral*, l'autre frère, constitue le topique-propos (*aboutness topic*) de l'expression *isi saranda e brešengo*, a quarante ans. Les deux phrases sont symétriques.

Interprétation sémantique

La référence des deux syntagmes se calcule chaque fois par l'adjonction des sèmes de l'adjectif et de ceux du nom : le référent de l'expression globale est à l'intersection de l'ensemble {frères} et de l'ensemble {entités grandes} et {autres entités}. L'identification du référent se fait en opérant une restriction pragmatique au sein des référents du nom. Les deux adjectifs portent un sème spécifique : le référent de *o baro phral*, le grand frère, est contrastif du référent de *o vaver phral*, l'autre frère, mentionné dans la phrase suivante. Il existe donc un focus étroit à valeur contrastive portant sur les adjectifs *baro*, grand, et *vaver*, autre.

(191) – (192) M46A, Korçë, 24 juillet 2014

Les explications suivantes portent sur les différents groupes de Roms, et sont issues d'une conversation spontanée.

1	*M46A: ə/ jek fisi vaver/ kaj vaćeren ende Čergarja/ ale em Grećijata Čergarja/ a odova vakət.*
2	*Em/ em but xanik Čergarja avile/ Bitolatar/ Manastirendar.*
3	*O maj purane Roma kaj si andi Albanja isi o/ o Mečkarja kaj će isi/ pandž šel/ pandž šel breš/*
4	*pandž šelupenda breš va.*
5	*Akale si maj purane Roma.*
6	*O vaver Roma andi Albanja/ nanaj po but se šelubiš seutrijanda breš.*
1	M46A : euh/ un autre groupe/ qu'on appelle Čergars/ sont venus aussi de Grèce les Čergars/ ah à cette époque.
2	Et/ et très peu de Čergars sont venus/ de Bitola/ de Manastir.
3	Les plus vieux Roms qui sont en Albanie sont les/ les Mečkars qui sont là depuis/ cinq cents/
4	cinq cents ans/ cinq cent cinquante ans oui.
5	C'est eux les plus anciens Roms.
6	Les autres Roms en Albanie/ ne sont pas là depuis plus de cent vingt cent trente ans.

(191)	***O***	***maj***	***puran-e***	***Rom-a***	*kaj*	*si*
	ART.DEF.PL	COMP	ancien-PL	Rom-PL	REL	être.3.PRES
	andi	*Albanja*	*isi*	*o/*	*o*	*Mečkar-ja.*
	dans.ART.DEF.F.SG	Albania(F.SG)	être.3.PRES	ART.DEF.PL	ART.DEF.PL	Mečkar-M.PL

« Les plus vieux Roms qui sont en Albanie sont les/ les Mečkars […]. »

(192)	***O***	***vaver***	***Rom-a***	*andi*	*Albanja*	*nanaj*
	ART.DEF.PL	autre[PL]	Rom-PL	dans.ART.DEF.F.SG	Albania(F.SG)	être.NEG.3.PRES
	po	*but*	*se*	*šelubiš*	*seutrijanda*	*breš.*
	COMP	beaucoup	que	cent_vingt	cent_trente	année[M.PL]

« Les autres Roms en Albanie/ ne sont pas là depuis plus de cent vingt cent trente ans. »

Statut de l'information

Le référent de *o maj purane Roma*, les plus anciens Roms, est nouveau – même s'il a déjà été question d'autres groupes de *Roms*, Arli et Čergar. L'adjectif *purane* est nouveau lexicalement.

Le référent de *o vaver Roma*, les autres Roms, est accessible-agrégatif – il dénote un groupe constitué de plusieurs référents déjà mentionnés, Arli et Čergar. L'adjectif *vaver* est nouveau lexicalement.

Structure de l'information

Le syntagme *o maj purane Roma*, les plus anciens Roms, constitue le topique-propos de l'expression *isi o/ o Mečkarja*, sont les/ les Mečkars.

Le syntagme *o vaver Roma*, les autres Roms, constitue le topique-propos de l'expression *nanaj po but se šelubiš seutrijanda breš*, ne sont pas là depuis plus de cent vingt cent trente ans.

Interprétation sémantique

La référence se calcule chaque fois par l'adjonction des sèmes de l'adjectif et de ceux du nom : le référent de l'expression globale est à l'intersection de l'ensemble {Roms} et de l'ensemble {entités [les plus] anciennes} ou {autres entités}. L'identification du référent se fait en opérant une restriction pragmatique au sein des référents du nom. Le référent *o maj purane Roma*, les plus anciens Roms, est contrastif du référent *o vaver Roma*, les autres Roms, mentionné deux phrases plus loin. Il existe donc un focus étroit à valeur contrastive portant sur les adjectifs *purane*, anciens, et *vaver*, autre.

2.9. Corpus spontané : DEF N/ DEF A

(194) F60A, Korçë, 21 juillet 2014

Dans l'entretien dirigé suivant, la locutrice présente sa famille.

1 *F60A: Ti organi/ si la organizàta kaj euh dikhèla e čorrorenge, dikhèla akale maksemènge.*
2 *Oj ə/ akale Romenge butjaće prastàla, e Romenge prastala.*
3 *M46A: Pomožinel le/*
4 *F60A: Pomožinèla e romenge/ zèma/*
5 *Akəja tradita denas laće mo dad miro.*
6 *Astarèla isine pala pèste/ kan isine mo dad presidènti/ e romèngoro.*
7 *Em denas la ov akija/ akova drom e butjakoro.*
8 *Hem oj akova o drom akana/ akale butja sa lela pes.*
9 *Em isi laće/ me dadestar/ akja buti.*
10 *M46A: OK. Euh/ tu/ keti škola isi tut?*
11 *F60A: Me na gelom but ti škola soske barjardom me phenje/ barjardom me phrales/ e xurdes/ e Benes em e Moza barjardom/ em nikadèma kati škola.*
12 *M46A: Oh/*
13 *F60A: Eh/ so ka keras? Akhal isine o breša o zaman.*
14 *M46A: Na/ džanes soske pučhlom?*
15 *Akaleske: sar kerdan trektìja tu?*
16 *Pindžarèsa/ avèsa o love/*
17 *F60A: Me o love ah/ sa/*
18 *M46A: pindžarèsa sa/*
19 *F60A: I školìa but nanaj man por me te kerav o/ melònja/ džanàva te ginav.*
20 *Em džanàv te kerav buti but šukar.*
21 *M46A:* [rire] *Brava!*

1	F60A : Dans l'organi/ elle a une organisation qui euh s'occupe des pauvres, elle s'occupe de ces enfants.
2	Elle euh/ elle travaille pour ces Roms, elle travaille pour les Roms.
3	M46A : Elle aide les/
4	F60A : Elle aide les Roms, (?)/
5	Cette tradition c'est mon père à moi qui lui a donnée.
6	Il prenait après lui/ quand mon père il était président/ des Roms.
7	Et il lui a donné, lui, cette/ cette façon de travailler.
8	Et elle cette façon maintenant, ces choses elle a tout pris pour elle.
9	Et c'est pour elle/ ce travail/ de mon père.
10	M46A : OK. Euh/ toi/ combien de temps tu as été à l'école ?
11	F60A : Moi j'ai pas été beaucoup à l'école parce que j'ai élevé mes sœurs/ j'ai élevé mon frère/ le petit/ Beni et Moza j'ai élevé/ et je suis partie de l'école.
12	M46A : Oh/
13	F60A : Eh/ qu'est-ce qu'il fallait faire ? C'était comme ça à l'époque$_{romani}$ à l'époque$_{turc}$.
14	M46A : Non/ tu sais pourquoi je demande ?
15	Pour ça : comment tu as fait les tractations toi ?
16	Tu connais/ comment vient l'argent/
17	F60A : Moi l'argent ah/ tout/
18	M46A : Tu connais tout/
19	F60A : Je ne suis pas beaucoup instruite, mais moi pour faire les/ melons/ je sais compter.
20	Et je sais très bien travailler.
21	M46A : [rire] Bravo !

(194)	*Barjar-d-om*	***m-e***	***phral-es/***	***e***	***xurd-es/***	*e*
	élever-PRET-1SG	POSS1SG- OBL	frère-M.SG.OBL	ART.DEF.OBL	tout_petit-M.SG.OBL	ART.DEF.OBL
	Ben-es	*em*	*e*	*Moza*	*barjar-d-om.*	
	Beni-M.SG.OBL	et	ART.DEF.OBL	Moza(F.SG)	élever-PRET-1SG	

« J'ai élevé mon frère/ le tout petit/ Beni et Moza j'ai élevé. »

Le syntagme présente une première pause de 0,35 seconde puis une deuxième de 0,36 seconde.

Statut de l'information

Dans cet exemple, le référent est donné-inactif car la locutrice a déjà listé, plus haut dans l'entretien (conversation de l'exemple (154), l'ensemble de ses frères et sœurs. Tout au plus peut-on arguer qu'ils font maintenant partie de la connaissance du monde de l'interlocuteur, donc sont accessibles. Le nom et l'adjectif sont nouveaux lexicalement dans cette partie de l'entretien.

Structure de l'information

Le syntagme *me phrales/ e xurdes/ e Benes*, mon frère/ le petit/ Beni, constitue le focus de nouvelle information sollicitée avec le verbe *barjardom*, j'ai élevé. Si une alternative était explicitement donnée en discours, on aurait en fait un focus de contraste – *subtype selecting* (Götze et al. 2007).

Interprétation sémantique

La partie nominale du syntagme, le frère, fait partie du contexte, du présupposé, puisque la locutrice a déjà présenté ses frères et sœurs. Le contraste se fait donc *in presentia*. Elle vise, par ce syntagme, à sélectionner l'un d'entre eux : le plus petit. La référence se calcule par l'adjonction des sèmes de l'adjectif et de ceux du nom. En d'autres termes, le référent de l'expression globale est à l'intersection de l'ensemble {frères} et de l'ensemble {toutes petites entités}. L'identification du référent se fait en opérant une restriction pragmatique au sein des référents du nom grâce à l'adjectif. L'adjonction du nom propre permettant de mieux identifier le référent constitue un exemple d'apposition nominale, et rappelle l'apposition de Tiranë dans l'exemple (185). Il s'agit d'une modification indirecte.

(197) – (200) F47M, Llakatund, 13 août 2014

Dans l'entretien dirigé suivant, la locutrice présente sa famille.

1	*F47M: Si man trin. Duj čhaja ta jek čhaʊo. Si man.*
2	*O čhaja si prandime, me čhaja.*
3	*Ch:* [question hors micro]
4	*F47M: Ah/ Temin. O čhija miɻe si pɻandime.*
5	*Jik isi Tiɻan. Thela duje tikinen.*
6	*I li odja ave si Beɻat. I li oj duje tikinen thela.*
7	*O čhaʊo si ki škola, pəɻmieksi. Hmm.*
8	*Ch:* [question hors micro]
9	*F47M: I čhaj/ bučhola Nəɻtila i baɻi. Nərtila.*
10	*I djetyto bučhola Brikena.*
11	*O čhaʊo bučhola Mikele. Hmm.*
12	*Ch:* [question hors micro]
13	*F47M: Isi man phɻala isi ma duj phɻala isinam ə/ pandž phenja. Eh/*
14	*Ch:* [question hors micro]
15	*F47M: ə/ i baɻi phen miɻi bučhos la Lila.*
16	*Odəja aver Faɻija. Špɻesa. Isinom ə Drita. I Valide ə po tikini.*
17	*Me phrala isi o Astriti, o Gimi. ə/*
18	*Ch:* [question hors micro]
19	*F47M: Akoʊa bešen Elbasan. ðeðe me isinom akate Vloɻ. Ah.*
1	F47M : J'en ai trois. Deux filles et un garçon. Oui j'en ai.
2	Les filles sont mariées, mes filles.
3	Ch : [question hors micro]
4	F47M : Ah/ (?). Mes filles sont mariées.
5	Une est à Tiranë. Elle a deux petits.
6	La grande aussi est à Berat. Elle aussi elle a deux petits.
7	Le garçon va à l'école, (?). Hmm.
8	Ch : [question hors micro]
9	F47M : La fille s'appelle Nertila la grande. Nertila.
10	La deuxième s'appelle Brikena.
11	Le garçon s'appelle Mikele. Hmm.
12	Ch : [question hors micro]

13	F47M : J’ai deux frères j’ai deux frères on a euh/ cinq sœurs. Eh/
14	Ch : [question hors micro]
15	F47M : Euh/ ma grande soeur s’appelle Lila.
16	L’autre s’appelle Fariha. Shpresa. Moi je suis Drita. Valide la plus petite.
17	Mes frères c’est Astrit, Gimi. Euh/
18	Ch : [question hors micro]
19	F47M : Eux ils habitent à Elbasan. Il y a que moi qui suis ici à Vlorë. Ah.

(197)	***I***	***čhaj/***	*buč-ola*	*Nəɻtila*	***i***	***baɻ-i***.
	ART.DEF.F.SG	fille(F.SG)	s’appeler-3SG.PRES	Nertila(F.SG)	ART.DEF.F.SG	grand-F.SG

« La fille s’appelle Nertila la grande. »

Le syntagme présente une pause de 0,24 seconde suivie d’une insertion.

Statut de l’information

Le référent est donné-inactif car la locutrice a déjà mentionné ses enfants. Le nom est donné-inactif, mais l’adjectif est nouveau lexicalement.

Structure de l’information

Le syntagme *i čhaj* […] *i bari*, la grande […] fille, constitue le topique-propos de *bučhola Nəɻtila*, s’appelle Nertila.

Interprétation sémantique

La partie nominale du syntagme, la fille, fait partie du présupposé. Le syntagme vise à sélectionner deux référents au sein de l’ensemble {filles}, grâce à l’adjectif *grande*. Le référent global est donc à l’intersection de ceux du nom et de ceux de l’adjectif. Il s’agit d’un syntagme polydéfini à effet de sélection classique où l’adjectif est un modifieur indirect. La rupture se manifeste par une pause et par l’insertion de tout le syntagme verbal, *bučhola Nəɻtila*, s’appelle Nertila. La « coupure » du syntagme nominal en deux est selon nous possible car le syntagme verbal inséré n’est pas long : sujet + complément, mais pas d’autre élément.

(200)	*I li*	***odja***	***ave***	*si*	*Beɻat.*
	FOC	DEM.F.SG	autre	être.3PRES	Berat

« L’autre aussi est à Berat. »

Statut de l’information

Le référent est donné-inactif car la locutrice a déjà mentionné ses enfants. L’adjectif est lexicalement nouveau.

Structure de l'information

Le syntagme *odja ave*, l'autre, constitue le topique-propos de *si Beɻat*, est à Berat.

Interprétation sémantique

Ce syntagme comprend seulement une partie adjectivale, où l'adjectif est sous focus étroit. Ce focus est contrastif, puisqu'il oppose une fille à une autre, au moyen de la propriété adjectivale *autre*. Le référent se trouve donc à l'intersection de la dénotation du nom *fille*, qui est sous-entendu mais présent un peu plus haut dans la conversation, et de la dénotation de l'adjectif *autre*. Le lien avec le nom-antécédent se fait au moyen du démonstratif féminin *odja*. Le focus est en outre exprimé par la particule de focalisation *i li*, aussi – ce qui n'empêche pas l'ensemble du syntagme *i li odja ave*, l'autre aussi, de jouer le rôle de topique-propos de l'énoncé.

2.10. Corpus spontané : DEF A et DEM N

(198) M58MČ, Voskpojë, 15 juillet 2013

L'échange suivant est issu d'une conversation spontanée.

1	*M58MČ: Kazom mase istarel ə/ e/ džuvli/ romni/ e maksemi ano por?*
2	*Ch: A/ enja/ mesek.*
3	*M58MČ: Enja.* [phrase en albanais] *ə/ s/*
4	*Kana e drija šunel e maksemen? Kaj si ano por?*
5	*Ch: ə/*
6	*M58MČ: So m/ so mase?*
7	*Ch: e dəj sam/*
8	*M58MČ: Sabu?*
9	*Ch: So si ə/ so puče?*
10	*M58MČ: Kana e dej/ ašundel e makseme/ e/ e cekne ano por?*
11	*So mase ašunel? Na/ ni džane? ko dujto mase.*
12	*Vov/ čheljal.*
13	*Ch: Aha. <ćhelel.>*
14	*M58MČ: <Ov> ćheljal. Aha.*
15	*Ch: Ando por.* [laughs]
16	*M58MČ: Ano por, va.*
17	*Va, ćhelel vi vov va.*
1	M58MČ : Combien de mois elle garde euh/ la/ femme/ la femme rom/ l'enfant dans le ventre ?
2	Ch : Ah/ neuf/ mois.
3	M58MČ : Neuf. [phrase en albanais] euh/ s/
4	Quand la (?) entend-elle l'enfant ? qui est dans le ventre ?
5	Ch : euh/
6	M58MČ : Quel m/ quel mois ?
7	Ch : la mère seulem/

8	M58MČ : (?)
9	Ch : Qu'est-ce que c'est euh/ qu'est-ce que tu demandes ?
10	M58MČ : Quand est-ce que la mère/ entend l'enfant/ le/ le petit dans le ventre ?
11	Quel mois elle entend ? Non/ tu sais pas ? au deuxième mois.
12	Il danse.
13	Ch : Aha. <Il danse.>
14	M58MČ : <Lui> il danse. Aha.
15	Ch : Dans le ventre. [rire]
16	M58MČ : Dans le ventre, oui
17	Oui, il danse lui aussi oui.

(198)	*Kana*	*e*	*dej/*	*ašun-d-el*	***e***	***maksem-e/***
	quand	ART.DEF.F.SG.OBL	fille(F.SG)	écouter-PRET-3.PRES	ART.DEF.M.SG.OBL	enfant-M.SG.OBL
	e/	***e***	***cekn-e***	*an-o*	*por?*	
	ART.DEF.M.SG.OBL	ART.DEF.M.SG.OBL	petit-M.SG.OBL	dans-ART.DEF.M.SG	ventre(M.SG)	

« Quand est-ce que la fille/ entend l'enfant/ le/ le petit dans le ventre ? »

Le syntagme présente une pause de 0,29 seconde.

Statut de l'information

Le référent est donné-inactif car il en a déjà été question deux fois. Le nom est donné-inactif, mais l'adjectif est nouveau lexicalement.

Structure de l'information

Le syntagme *e makseme/ e/ e cekne*, l'enfant/ le/ le petit, constitue le focus de nouvelle information sollicitée, avec le verbe *ašundel*, entend.

Interprétation sémantique

La partie nominale du syntagme, l'enfant, fait partie du contexte, du présupposé. Cependant l'adjectif ne vise pas, selon nous, à sélectionner un petit enfant parmi d'autres. En d'autres termes, le référent de l'expression globale n'est pas à l'intersection de {enfants} et de {entités petites}, car ces deux ensembles sont identiques dans ce contexte. Il s'agit d'un fœtus dans le ventre de la mère, donc, aux yeux du locuteur, d'un « enfant » qui est « petit ». En outre, le référent de « le petit » a un sens générique. On pourrait considérer que l'adjectif transmet une propriété inhérente (« le petit enfant » pour « le fœtus »), mais cela serait contradictoire avec son emplacement postposé, qui fait de lui un modifieur indirect. Pour cette raison, nous considérons qu'il ne s'agit pas d'un syntagme polydéfini, mais de deux syntagmes nominaux parallèles, où le second comprend un adjectif substantivé.

(199) M22A, Korçë, 6 août 2014

L'extrait suivant est issu d'un entretien dirigé où le locuteur présente sa famille.

1	*Ch: Phrala phenja isi tut?*
2	*M22A: Daa jek khorak?*
3	*Ch: Phenja phrala isi tut?*
4	*M22A: Isi man duj phenja.*
5	*Ch: Bare/ po bare?*
6	*M22A: Jek isi po bari manda jek isi po tikni.*
7	*Ch: Prandime isi?*
8	*M22A: Odəja bari isi prandime.*
9	*I tikni isi bešar.*
1	Ch : Tu as des frères et sœurs ?
2	M22A : Est-ce que tu peux répéter ? (litt. encore une fois ?)
3	Ch : Tu as des sœurs et frères ?
4	M22A : J'ai deux sœurs.
5	Ch : Grandes/ plus grandes ?
6	M22A : Une est plus grande que moi une est plus petite.
7	Ch : Elles sont mariées ?
8	M22A : La grande est mariée.
9	La petite est célibataire.

(199)	***Odəja***	***bar-i***	*isi*	*prandime.*
	DEM.F.SG	grand-F.SG	être.3PRES	marié

« La grande est mariée. »

Statut de l'information

Le référent est donné-inactif car le locuteur a déjà mentionné ses frères et sœurs – ou donné-actif si l'on prend seulement en compte les tours de parole du locuteur. L'adjectif est lexicalement donné, actif ou inactif pour la même raison.

Structure de l'information

Le syntagme *odəja bari*, la grande, constitue le topique-propos de *isi prandime*, est mariée.

Interprétation sémantique

Ce syntagme comprend seulement une « partie adjectivale », où l'adjectif est sous focus étroit. Il s'agit d'un focus contrastif, puisqu'il oppose une sœur à une autre, au moyen de la propriété adjectivale *grande*. Le référent se trouve donc à l'intersection de la dénotation du nom *sœur*, qui est sous-entendu mais présent un peu plus haut dans la conversation, et de la dénotation de l'adjectif *grande*. Le lien avec le nom-antécédent se fait au moyen du démonstratif féminin *odəja*.

2.11. Corpus spontané : DEF N/ A

(202) M71A, Korçë, 28 juillet 2014

L'extrait suivant est issu d'un récit de vie qui a suivi l'entretien dirigé où le locuteur se présentait.

1	*Ch: Tu dikhlan akala?*
2	*M71A: Ka/ džən/ dikhlom me! Sa ka vaćerenas amenge/*
3	*Po than gelam ote amen/ kaj gelam/ me somas kaing biše brešengo.*
4	*Dikhas sine o bumbe akhal sar aljena/ o pištolja/ purane kaj phablea isine/ kupe/ kaj ačhiljas e Grećešće.*
5	*Dikhlom le me jaćhensa odole.*
1	Ch : Toi tu les as vus ?
2	M71A : (?)/ hein/ j'ai vu moi ! Qu'est-ce qu'ils nous racontaient/
3	On y a été sur place nous/ qu'on y a été/ moi j'avais à peu près vingt ans.
4	On voyait les bombes comme ça qui arrivaient/ les pistolets/ anciens qui tiraient/ (?)/ qui restaient des Grecs.
5	Je l'ai vu de mes propres yeux ça.

(202)	*Dikh-as*	*sine*	*o*	*bumb-e*	*akhal*	*sar*	*alj-ena/*
	voir-1PL	être.3.PRET	ART.DEF.PL	bombe-F.PL	comme_ça	comme	venir.PRET-3PL
	o	***pištol-ja/***	***puran-e***	*kaj*	*phab-lea*	*isine*	[…].
	ART.DEF.PL	pistolet-M.PL	ancien-PL	REL	brûler-3SG	être.3.PRET	

« On voyait les bombes comme ça qui arrivaient/ les pistolets/ anciens qui tiraient […]. »

Le syntagme présente une pause de 0,2 seconde.

Statut de l'information

Le référent est nouveau – mais on pourrait arguer qu'il fait partie du champ lexical de la guerre, avec *bombe*, bombes, et serait donc accessible-inférable. Le nom comme l'adjectif sont nouveaux lexicalement.

Structure de l'information

Le syntagme *o pištolja/ purane*, les pistolets/ anciens, constitue le focus de nouvelle information non sollicitée.

Interprétation sémantique

La référence se calcule par l'adjonction des sèmes de l'adjectif et de ceux du nom : le référent de l'expression globale est à l'intersection de l'ensemble {pistolet} et de l'ensemble {entités ancienne}. On peut hésiter quant à la valeur du focus étroit : il peut avoir une valeur contrastive portant sur l'adjectif, qui l'oppose à l'adjectif « modernes », ou bien une valeur assertive, emphatisant seulement

le sens de l'adjectif. Le syntagme en (202) est un exemple de modification indirecte. Le contraste se fait *in absentia*.

(203) F22M, Llakatund, 13 août 2014

Dans l'entretien dirigé suivant, la locutrice parle de tradition religieuse.

1	*Ch: I tradita andi familja i/ religja sar isi?*
2	*F22M: ə/ mi familja/ respektonela i tradita/ ə ɻomani/*
3	*ə/ Herdelezi çdo bersh kaj si, amen kerasa amare manušensa e khelimasa jekhe khereste hem bešasa derjavin.*
4	*ə/ o urjajpa na but.*
5	*Hmm/ nane as mi daj as mo dad ə/ as me manuša kaj/ thena o uɻjajpa ɻoman/ ə saɻ o ɻoma urjardjona pomen ə/ i čhib theɻaha la ə/ ta/ akala.*
1	Ch : La tradition dans la famille la/ religion c'est comment ?
2	F22M : Euh/ ma famille/ respecte la tradition/ euh rom/
3	euh/ Herdelezi qui a lieu chaque année, nous on le fait avec les gens de ma famille avec les danses dans une maison et on va dans le fleuve.
4	Euh/ les vêtements pas beaucoup.
5	Hmm/ il y a personne ni ma mère ni mon père euh/ ni les gens de ma famille qui/ mettent les vêtements rom/ euh comme les Roms s'habillent euh/ la langue on la garde euh/ et/ voilà.

(203)	*ə/*	*m-i*	*familja/*	*respekton-ela*	***i***	***tradita/***	***ə***	***ɻoman-i***	[…]
	3PL	POSS.1SG-F.SG	famille(F.SG)	respecter-3SG.PRES	ART.DEF.F.SG	tradition(F.SG)	euh	rom-F.SG	

« Euh/ ma famille/ respecte la tradition euh/ rom […]. »

Le syntagme présente une pause de 0,64 seconde suivie d'une insertion entre le nom et l'adjectif.

Statut de l'information

Le référent est accessible-inférable de la question qui vient d'être posée. Le nom est donc donné-accessible, mais l'adjectif est nouveau lexicalement.

Structure de l'information

Le syntagme *i tradita ə/ ɻomani*, la tradition euh/ rom, constitue le focus de nouvelle information sollicitée, avec le verbe *respektonela*, respecte.

Interprétation sémantique

La référence se calcule par l'adjonction des sèmes de l'adjectif et de ceux du nom : le référent de l'expression globale est à l'intersection de l'ensemble {tradition} et de l'ensemble {entités roms}. Il

existe donc un focus étroit à valeur contrastive portant sur l'adjectif, qui l'oppose à l'adjectif *gadžikani*, non-rom. Le syntagme en (203) est un exemple de modification indirecte. Le contraste se fait *in absentia*.

(204) M46A, Korçë, 24 juillet 2014

L'extrait suivant est issu d'une conversation spontanée.

1	*M46A: Maj but o Roma/ maj but Roma a/ avile Grećijata.*
2	*But xanik Roma avile bə/ Grećijata.*
3	*Kana ale ə/ o Čamja/ kan ani Grećija tandas e čamen.*
4	*Em odole Roma ka ale Čamjensar isi/ kaj bešena/ samo/ Fjeri.*
5	*Em olenge vaćerena/ Bamilja/ Bamilja.*
6	*Ole džanena/ sa i/ i čhib ə/ gadžikani grećist.*
1	M46A : La plupart des Roms/ la plupart des Roms s/ sont venus de Grèce.
2	Très peu de Roms sont venus euh/ de Grèce.
3	Quand sont venus euh/ les Čams/ quand en Grèce on a chassé les Čams.
4	Et ces Roms qui sont arrivées avec les Čams sont/ ils habitent/ seulement à Fieri.
5	Et on les appelle/ Bamilja/ Bamilja.
6	Eux ils parlent tous la/ la langue euh/ non Rom grecque.

(204)	*Ole*	*džan-ena/*	*sa*	*i/*	***i***	***čhib***	***ə/***	***gadžikan-i***	*grećist*
	3PL	savoir-3PL.PRES	tout	ART.DEF.F.SG	ART.DEF.F.SG	langue(F.SG)	euh	non_rom-F.SG	en_grec

« Eux ils parlent tous la/ la langue euh/ non Rom grecque. »

L'exemple présente une insertion entre le nom et l'adjectif suivie d'une pause de 0,38 seconde.

Statut de l'information

Le référent est nouveau. Le nom est nouveau lexicalement, le premier adjectif (*gadžikani*, non-rom) est nouveau et le deuxième adjectif (*grećist*) donné-inactif lexicalement. *Grećist* est en principe un adverbe, mais il trouve ici un emploi adjectival du fait de sa juxtaposition à *gadžikani*.

Structure de l'information

Le syntagme *i čhib ə/ gadžikani grećist*, la langue euh/ non Rom grecque, constitue le focus de nouvelle information sollicitée, avec le verbe *džanena*, ils parlent.

Interprétation sémantique

La référence se calcule par l'adjonction des sèmes de l'adjectif et de ceux du nom : le référent de l'expression globale est à l'intersection de l'ensemble {langue} et de l'ensemble {entités non-roms}.

Il existe donc un focus étroit à valeur contrastive portant sur l'adjectif, qui l'oppose à l'adjectif *romani*, rom : on peut dire qu'il y a contraste *in presentia*. Le syntagme en (204) est un exemple de modification indirecte. La juxtaposition de l'adverbe *grećist*, en grec, dans le rôle d'adjectif, vise à mieux identifier le référent exact de « langue non-rom ». Cette tactique est la même que dans les exemples (185) et (194).

3. Annexes du chapitre V : Analyse du corpus semi-spontané

3.1 Métadonnées et codage

Tableau 121 : Les couples de participants du test 1

Participant A	Participant B	Durée de l'enreg. (mn:sec)
F16AČ	la chercheuse	4:49
M46A	la chercheuse	6:28
F16AČ	M60A	1:46
M60A	F16AČ	6:34
F11A	F16AČ	1:57
F9A	F16AČ	2:28
F16AČ	F38A	3:35
F38A	F16AČ	1:41
M13AČ	F27A	3:37
F20A	M13AČ	2:11
F22M	F60A	3:14
F47M	F22M	5:10
M19M$_{(2)}$	F22M	4:00
M22a	M45M	5:17
M45M	M22a	7:10
F24AČ	F16AČ	1:35

Tableau 122 : Les couples de participants du test 2

Participant A	Participant B	Durée de l'enreg. (mn:sec)
F16AČ	M13AČ	1:18
M13AČ	F16AČ	1:34
F16AČ	M46A	2:02
F16AČ	M46A	4:26
M46A	F16AČ	3:02
M46A	M43X	3:48
M43X	M46A	5:16
M37A	M34AČ	3:18
M34AČ	M37A	2:17
F16AČ	F60A	0:53
F16AČ	F60A	3:10
F16AČ	F60A	2:23
F60A	F16AČ	2:00
F60A	F16AČ	1:41
M13AČ	F22M	2:24

Participant A	Participant B	Durée de l'enreg. (mn:sec)
F22M	M13AČ	1:48
F16AČ	F27A	1:39
F27A	F16AČ	1:22
M13AČ	F44AČ	2:38
F44AČ	M13AČ	6:45
F44AČ	F22AČ	3:50
F22AČ	F44AČ	3:11
F85A	F44AČ	1:00[1]
		0:30
		0:04
		0:11
		0:17
		0:26
		0:01
		0:50
F16AČ	M18Č	2:35
M18Č	F16AČ	2:20
F44AČ	F53A	4:02
F53A	M52A	3:47
M52A	F53A	3:28
F37MB	M58Č	5:05
M58Č	F37MB	4:09
F66M	M58MČ	2:42
M58MČ	F66M	5:09
M58MČ	M19M(1)	5:09
M19M(1)	M18Č	3:27
M18Č	F23AČ	2:56
F23AČ	M18Č	2:34
F44AČ	M71A	4:32
M71A	F44AČ	6:06
F24K	M27Dž	4:46
M46A	M27Dž	6:01
M27Dž	M46A	5:32
F65A	F85A	2:25
F65A	F85A	5:15
F65A	F85A	4:21
F85A	F65A	4:23
F16AČ	F38A	4:03
F38A	F16AČ	3:08
M13AČ	F27A	5:36
F27A	M13AČ	4:26

[1] L'enregistrement a été entrecoupé, en raison d'un problème de batterie de l'enregistreur.

Participant A	Participant B	Durée de l'enreg. (mn:sec)
M24	M22a	5:41
M22a	M24	9:53
F22M	M19M$_{(2)}$	10:02
M19M$_{(2)}$	F22M	10:25
F47M	M19M$_{(2)}$	6:57
M45M	M22a	0:48
		7:41
F24AČ	F16AČ	4:50

Tableau 123 : Profil des locuteurs[1] ayant participé aux tests 1 et 2

Nom	Age	CLASSE D'AGE	GENRE	Lieu d'enreg.	Lieu de naissance / de résidence[2]	VARIETE	Autres langues connues[3]	EDUCATION	Profession	MILITANTISME	Tests réalisés	Temps[4]
F9A	9	0	1	Bilisht	Bilisht	2	albanais	0	écolière	0	1	1 / 2:28
M11A	11	0	0	Korçë	Korçë	2	albanais	0	écolier	0	2	1 / 1:48
F11A	11	0	1	Bilisht	Bilisht	2	albanais	1	écolière	0	1	1 / 1:57
M13AČ	13	0	0	Korçë	Korçë	4	albanais	1	écolier	0	1 & 2	1 / 3:37 4 / 12:12
F16AČ	16	1	1	Korçë	Korçë	4	albanais	1	écolière	0	1 & 2	3 / 10:10 9 / 22:29
M18Č	18	1	0	Korçë, Voskopojë	Fier	3	albanais (+ anglais)	1	écolier	0	2	2 / 5:16
M19M(1)	19	1	0	Voskopojë	Xhafzotaj	1	albanais, italien, anglais	2	étudiant	0	2	1 / 3:27
M19M(2)	19	1	0	Llakatund	Llakatund / Fier	1	albanais (+ italien, anglais)	2	écolier	0	1 & 2	1 / 4:00 1 / 10:25
F20A	20	1	1	Korçë	Korçë / Masa	2	albanais, italien	1	cuisinière, puis sans emploi	0	1 & 2	1 / 2:11 1 / 4:26
F22M	22	1	1	Llakatund	Tirana	1	albanais (+ anglais, italien)	2	étudiante	1	1 & 2	1 / 3:14 1 / 10:02
M22a	22	1	0	Berat	Korçë / Tiranë	0	albanais, français, anglais, italien, espagnol	2	journaliste, puis employé au ministère des affaires sociales	1	1 & 2	1 / 5:17 1 / 9:53
F22AČ	22	1	1	Korçë	Korçë	4	albanais (+ italien)	1	journaliste	0	2	1 / 3:11

[1] Les locuteurs témoins sont en italique.

[2] Nous mentionnons derrière la barre oblique le lieu de résidence si celui-ci est ou a été différent du lieu de naissance.

[3] La mention des langues connues autres que le romani correspond aux pratiques déclarées des locuteurs. J'ai distingué deux niveaux de connaissance : plutôt fluide ou bien plutôt partielle (indiquée entre parenthèses).

[4] Nombre d'enregistrements / durée en minutes:secondes

Nom	Age	CLASSE D'AGE	GENRE	Lieu d'enreg.	Lieu de naissance / de résidence[2]	VARIETE	Autres langues connues[3]	EDUCATION	Profession	MILITANTISME	Tests réalisés	Temps[4]
F23AČ	23	1	1	Voskopojë	Tiranë	4	albanais (+ espagnol)	0	travailleuse associative	0	2	1 / 2:34
F24K	*24*	*1*	*1*	*Ohrid*	*Tetovo*	*8*	*macédonien, anglais, serbe, croate (+ turc)*	*2*	*étudiante*	*0*	*2*	*1 / 4:46*
M24AB	24	1	0	Berat	Berat	5	albanais (+ grec)	1	enseignant de romani	1	2	1 / 5:41
F24AČ	24	1	1	Korçë	Korçë / Tiranë, Roma	4	albanais, italien	1	coiffeuse, puis sans emploi	0	1 & 2	1 / 1:35 1 / 4:50
F27A	27	1	1	Korçë	Korçë	2	albanais	0	vendeuse	0	2	1 / 1:22
M27Dž	*27*	*1*	*0*	*Ohrid*	*Tetovo*	*9*	*macédonien, albanais, turc, français*	*2*	*employé administration*	*0*	*2*	*1 / 5:32*
M34AČ	34	1	0	Korçë	Korçë / Florina	4	albanais, grec	0	ouvrier et divers emplois	0	2	1 / 2:17
M37A	37	1	0	Korçë	Korçë / Florina	2	albanais, grec	0	vendeur	0	2	1 /3:18
F37MB	37	1	1	Voskopojë	Tiranë /	7	albanais, anglais, italien	2	juriste	1	2	1 / 5:05
F38A	38	1	1	Korçë	Pogradec / Korçë, Masa	2	albanais, italien	1	cuisinière, puis sans emploi	0	1 & 2	1 / 1:41 1 / 3:08
M43X	*43*	*2*	*0*	*Pogradec, Korçë*	*Vuçitërn[?], Kosovo / Kassel*	*10*	*serbe, allemand, albanais, anglais, français, « Broken English », italien, polonais, russe (+macédonien, turc, arabe)*	*1*	*homme d'affaires*	*0*	*2*	*1 / 5:16*
F44AČ	44	2	1	Korçë	Fier / Korçë	4	albanais, grec	0	femme au foyer	0	2	4 / 19:09
M45M	45	2	0	Llakatund	Vlorë / Llakatund	1	albanais, italien, grec	2	officier, puis sans emploi, puis médiateur	0	1 & 2	1 / 7:10 2 / 8/29

Nom	Age	CLASSE D'AGE	GENRE	Lieu d'enreg.	Lieu de naissance / de résidence[2]	VARIETE	Autres langues connues[3]	EDUCATION	Profession	MILITANTISME	Tests réalisés	Temps[4]
M46A	46	2	0	Korçë	Korçë	2	albanais (+ grec)	1	journaliste	1	1 & 2	1 / 6:28 3 / 12:51
F47M	47	2	1	Llakatund	Elbasan / Llakatund	1	albanais	0	femme au foyer	0	1 & 2	1 / 5:10 1 / 6:57
M52A	52	2	0	Korçë	Korçë	2	albanais, grec, turc	1	ouvrier	0	2	1 / 3:28
F53A	53	2	1	Korçë	Korçë	2	albanais, grec (+ italien)	1	agricultrice puis femme au foyer	0	2	1 / 3:47
M58Č	58	2	0	Voskopojë	Lezhë / Mamurras	3	albanais	1	musicien à la retraite	0	2	1 / 4:09
M59MČ	59	2	0	Voskopojë	Tiranë	6	albanais, (+ serbe)	2	chorégraphe, enseignant de langue et culture romani, à la retraite	1	2	2 / 10:18
M60A	60	3	0	Bilisht	Bilisht	2	albanais, grec	0	vendeur	0	1	1 / 6:34
F60A	60	3	1	Korçë	Korçë	2	albanais (+ grec, + turc)	0	vendeuse à la retraite	0	2	2 / 3:41
F65A	65	3	1	Korçë	Korçë / Florina	2	albanais, grec	0	agricultrice puis femme au foyer et retraitée	0	2	3 / 12:01
F66M	66	3	1	Voskopojë	Fier / Vlorë, Korçë, Berat, Tiranë	1	albanais	1	enseignante de romani et militante	1	2	1 / 2:42
M71A	71	3	0	Korçë	Rembesi / Florina	2	albanais, grec (+ macédonien)	0	maçon puis vendeur, à la retraite	0	2	1 / 6:06
F85A	85	3	1	Korçë	Bilisht / Korçë	2	albanais (+ bulgaro-macédonien)	0	femme au foyer à la retraite	0	2	9 / 7:42

Tableau 124 : Codage des données

Type d'information	Information	Modalités possibles	Exemple	Remarques
Enregistrement	numéro de l'enregistrement	choix ouvert	12	
	nom du fichier audio	choix ouvert	Test 21 korrik - F16AČ à F60A ter.wav	
	date de l'enregistrement	choix ouvert	21 korrik 2014	
Profil locuteur	nom	choix ouvert	F16AČ	
	genre	0 = masculin 1 = féminin	1	
	âge numérique	choix ouvert	16	
	classe d'âge	0 – moins 15 ans 1 – 15 – 39 ans 2 – 40 – 59 ans 3 – plus 60 ans	1	
	niveau de formation	0 – 0 à 5 années de scolarité 1 – 6 à 12 années de scolarité 2 – plus 12 années de scolarité	1	
	lieu de naissance	choix ouvert	Korçë	
	lieu de résidence	choix ouvert	Korçë	Si la personne a vécu dans plusieurs lieux successifs, ils sont listés par ordre chronologique.
	variété familiale	0 – néo-locuteur arli 1 – mečkar 2 – arli 3 – čergar I 4 – arli & čergar I / čergar I & arli 5 – arli & bamlija 6 – mečkar & čergar I 7 – mečkar & bamlija 8 – kovački	4	Lorsqu'une variété est déclarée moins connue qu'une autre, elle est mise entre parenthèses.

Type d'information	Information	Modalités possibles	Exemple	Remarques
		9 – džambazo 10 – xoraxane		
	variété conjoint	0 – néo-locuteur arli 1 – mečkar 2 – arli 3 – čergar I 4 – arli & čergar I / čergar I & arli 5 – arli & bamlija 6 – mečkar & čergar I 7 – mečkar & bamlija 8 – kovački 9 – džambazo 10 – xoraxane	–	Si la personne n'est pas mariée : –
	autres langues connues	1 – mečkar 2 – arli 3 – čergar I 4 – arli & čergar I / čergar I & arli 9 – džambazo 11 – gadžikane : albanais, italien	albanais	Lorsqu'une langue est déclarée moins connue qu'une autre, elle est mise entre parenthèses.
SN prononcé	transcription du SN	choix ouvert	*jek tikno lapsi portokalija*	un petit crayon orange
Analyse	ordre des parties du discours	choix ouvert	IND A N A	
	déterminants	0 – article zéro 1 – DEF 2 – IND 3 – DEM 4 – POSS 5 – QUANT 6 – NUM 9 – plusieurs déterminants	2	
	nom	choix ouvert	*lapsi*	

Type d'information	Information	Modalités possibles	Exemple	Remarques
	adjectifs	choix ouvert	*tikno + portokalija*	
	type d'adjectif	0 – autres 1 – taille 2 – couleur 3 – aspect 9 – plusieurs types	9	La catégorie *autre* recouvre des adjectifs tels que *vaver*, autre, *džungalo*, dégoûtant, *desno*, de droite, *fundit*, dernier, *nevo*, nouveau…
	DEF N DEF A	0 = non 1 = oui	0	Chaque syntagme nominal est codé de manière binaire selon qu'il présente ou non la structure DEF N DEF A.
	DEF N A	0 = non 1 = oui	0	cf. le codage de DEF N DEF A.
	DEF A N	0 = non 1 = oui	0	cf. le codage de DEF N DEF A.
	complexes	0 = non 1 = oui	0	La catégorie *complexe* comprend les syntagmes mono- ou polydéfinis présentant plusieurs adjectifs (DEF A N A, DEF A N DEF A, DEF N DEF A DEF A…).
	DEM	0 = non 1 = pas d'A (N seul) 2 = pas de N (A seul) 3 = pas d'adjectif ni de nom 4 = A antéposé 5 = A postposé 9 = complexe	0	Si aucun DEM n'est présent dans le syntagme, il est codé par 0. Si un DEM est présent, le syntagme est codé par 1 si le démonstratif détermine un N seul, 2 s'il détermine un A seul, 3 s'il ne détermine ni un N ni un A (il est donc pronominalisé, ce qui n'implique pas l'absence de modifieurs). Si un DEM est présent et détermine à la fois un N et un A, on code 4 ou 5 selon la place de A par rapport au nom-tête, et enfin 9 dans le cas d'un syntagme nominal plus complexe (par exemple s'il y a à la fois un A antéposé et un A postposé ou s'il y a plusieurs démonstratifs).
	POSS	0 = non 1 = pas d'A (N seul) 4 = A antéposé 5 = A postposé	0	cf. le codage de DEM. Il n'existe pas, dans le corpus, de cas de POSS déterminant un A sans N (2), ni de possessif sans A ni N (3), et enfin pas de cas de structure complexe (9).
	IND	0 = non 1 = pas d'A (N seul) 2 = pas de N (A seul) 3 = pas d'adjectif ni de nom 4 = A antéposé 5 = A postposé 9 = complexe	9	cf. le codage de DEM.

Type d'information	Information	Modalités possibles	Exemple	Remarques
	NUM / QUANT	0 = non 1 = pas d'A (N seul) 3 = pas d'adjectif ni de nom 5 = A postposé 9 = complexe	0	cf. le codage de DEM. Il n'existe pas, dans le corpus, de cas de NUM ou de QUANT déterminant un A sans N (2), ni de syntagme où A soit antéposé au N (4).
	syntagme génitival	0 = non 1 = pas d'A (N seul) 2 = pas de N (A seul) 3 = pas d'adjectif ni de nom 4 = A antéposé 5 = A postposé 9 = complexe	0	La place de A est calculée par rapport au nom-tête, et non par rapport au nom du syntagme génitival, par exemple *i kapsa e gadengi zeleni*, la pince à linge verte, de structure DEF N DEF N=GEN A, est codée 5 car l'adjectif *zeleni*, verte, est postposé au nom-tête *kapsa*, pince à linge.
	proposition relative	0 = non 1 = pas d'A (N seul) 2 = pas de N (A seul) 3 = pas d'adjectif ni de nom 4 = A antéposé 5 = A postposé 9 = complexe	0	La place de A est calculée par rapport au nom-tête, et non par rapport à la proposition relative, par exemple *i xurdi kutulka kaj si čhinli*, la petite pomme de terre qui est coupée, de structure DEF A N SREL, est codée 4 car l'adjectif *xurdi*, coupée, est antéposé au nom-tête *kutulka*, pomme de terre.
	A ou N seul	0 = non 1 = pas d'A (N seul) 2 = pas de N (A seul) 3 = pas d'A ni de déterminant	0	cf. le codage de DEM.
	A et N seul	0 = non 4 = A antéposé 5 = A postposé 9 = complexe	0	cf. le codage de DEM.
	autre structure	0 = non 1 = oui	0	Lorsqu'en lieu et place d'un syntagme nominal, le locuteur prononçait autre chose (pronom, adverbe, proposition relative sans antécédent…), nous l'avons transcrit et codé *autre* (1).
	A antéposé	0 = non 1 = oui	1	Chaque syntagme nominal est codé de manière binaire selon qu'il présente ou non un A antéposé, quel que soit le déterminant présent.
	A postposé	0 = non 1 = oui	1	cf. le codage de A antéposé.

Ordre	Nom enregi	Locuteur	Genre	Age	Tranche d'âge	Lieu de naissance	Lieu de vie	Date enreg	Parler famille d'origine	Parler famille actuelle	Autres langues connues
27	26 korrik - Br à	Br	F	16	15 – 39 ans	Korça	Korça	26 korrik 2014	Arli + çergar	-	albanais
27	26 korrik - Br à	Br	F	16	15 – 39 ans	Korça	Korça	26 korrik 2014	Arli + çergar	-	albanais
27	26 korrik - Br à	Br	F	16	15 – 39 ans	Korça	Korça	26 korrik 2014	Arli + çergar	-	albanais
27	26 korrik - Br à	Br	F	16	15 – 39 ans	Korça	Korça	26 korrik 2014	Arli + çergar	-	albanais
27	26 korrik - Br à	Br	F	16	15 – 39 ans	Korça	Korça	26 korrik 2014	Arli + çergar	-	albanais
27	26 korrik - Br à	Br	F	16	15 – 39 ans	Korça	Korça	26 korrik 2014	Arli + çergar	-	albanais
27	26 korrik - Br à	Br	F	16	15 – 39 ans	Korça	Korça	26 korrik 2014	Arli + çergar	-	albanais
27	26 korrik - Br à	Br	F	16	15 – 39 ans	Korça	Korça	26 korrik 2014	Arli + çergar	-	albanais
27	26 korrik - Br à	Br	F	16	15 – 39 ans	Korça	Korça	26 korrik 2014	Arli + çergar	-	albanais
27	26 korrik - Br à	Br	F	16	15 – 39 ans	Korça	Korça	26 korrik 2014	Arli + çergar	-	albanais
27	26 korrik - Br à	Br	F	16	15 – 39 ans	Korça	Korça	26 korrik 2014	Arli + çergar	-	albanais
27	26 korrik - Br à	Br	F	16	15 – 39 ans	Korça	Korça	26 korrik 2014	Arli + çergar	-	albanais
27	26 korrik - Br à	Br	F	16	15 – 39 ans	Korça	Korça	26 korrik 2014	Arli + çergar	-	albanais
27	26 korrik - Br à	Br	F	16	15 – 39 ans	Korça	Korça	26 korrik 2014	Arli + çergar	-	albanais
27	26 korrik - Br à	Br	F	16	15 – 39 ans	Korça	Korça	26 korrik 2014	Arli + çergar	-	albanais
27	26 korrik - Br à	Br	F	16	15 – 39 ans	Korça	Korça	26 korrik 2014	Arli + çergar	-	albanais
27	26 korrik - Br à	Br	F	16	15 – 39 ans	Korça	Korça	26 korrik 2014	Arli + çergar	-	albanais
27	26 korrik - Br à	Br	F	16	15 – 39 ans	Korça	Korça	26 korrik 2014	Arli + çergar	-	albanais
27	26 korrik - Br à	Br	F	16	15 – 39 ans	Korça	Korça	26 korrik 2014	Arli + çergar	-	albanais
27	26 korrik - Br à	Br	F	16	15 – 39 ans	Korça	Korça	26 korrik 2014	Arli + çergar	-	albanais
27	26 korrik - Br à	Br	F	16	15 – 39 ans	Korça	Korça	26 korrik 2014	Arli + çergar	-	albanais
27	26 korrik - Br à	Br	F	16	15 – 39 ans	Korça	Korça	26 korrik 2014	Arli + çergar	-	albanais
27	26 korrik - Br à	Br	F	16	15 – 39 ans	Korça	Korça	26 korrik 2014	Arli + çergar	-	albanais
27	26 korrik - Br à	Br	F	16	15 – 39 ans	Korça	Korça	26 korrik 2014	Arli + çergar	-	albanais
27	26 korrik - Br à	Br	F	16	15 – 39 ans	Korça	Korça	26 korrik 2014	Arli + çergar	-	albanais
27	26 korrik - Br à	Br	F	16	15 – 39 ans	Korça	Korça	26 korrik 2014	Arli + çergar	-	albanais
27	26 korrik - Br à	Br	F	16	15 – 39 ans	Korça	Korça	26 korrik 2014	Arli + çergar	-	albanais
27	26 korrik - Br à	Br	F	16	15 – 39 ans	Korça	Korça	26 korrik 2014	Arli + çergar	-	albanais
27	26 korrik - Br à	Br	F	16	15 – 39 ans	Korça	Korça	26 korrik 2014	Arli + çergar	-	albanais
27	26 korrik - Br à	Br	F	16	15 – 39 ans	Korça	Korça	26 korrik 2014	Arli + çergar	-	albanais
27	26 korrik - Br à	Br	F	16	15 – 39 ans	Korça	Korça	26 korrik 2014	Arli + çergar	-	albanais
27	26 korrik - Br à	Br	F	16	15 – 39 ans	Korça	Korça	26 korrik 2014	Arli + çergar	-	albanais

Figure 91 : Codage du profil des locuteurs à chaque occurrence du corpus spontané

	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE	AF	AG
1	Déter	N	Type d'A	A	NP prononcé	Structure	DNDA	DNA	DAN	+ que D	Dem	Poss	Ind	Num	GenP	Relative	juste A c	juste AN	autre	A antép	A post
2	Art.Def	kultùlka	taille	bari	i kultùlka i bari	D N D A	oui	non	non	non	non	non	non	non	non	non	non	non	non	non	oui
3	Art.Def	kultùlka	taille	bari	pash t=i kultùlka t=i ba	P=D N P=D .	oui	non	non	non	non	non	non	non	non	non	non	non	non	non	oui
4	Art.Def	kultùlka	-	-	i kultùlka kaj si bangji	D N Rel	non	non	non	non	non	non	non	non	non	pas de A	pas de A	non	non	non	non
5	Art.Def	-	aspect	bangji	pash t=i bangji	P=D A	non	non	non	non	non	non	non	non	non	non	pas de N	non	non	non	non
6	Art.Def	-	taille	tikni	i tikni	D A	non	non	non	non	non	non	non	non	non	non	pas de N	non	non	non	non
7	Art.Def	kultùlka	taille	tikni	t=i tikni kultùlka opral	P=D A N Ad	non	non	oui	non	non	non	non	non	non	non	non	non	non	oui	non
8	Art.Indef	làpsi	-	-	jek làpsi kaj si portoka	Ind N Rel	non	non	non	non	non	non	pas d'A	non	non	pas de A	non	non	non	non	non
9	Art.Def	-	taille	tikni	i tikni	D A	non	non	non	non	non	non	non	non	non	non	pas de N	non	non	non	non
10	Art.Def	kultùlka	taille	tikni	pash t=i kultùlka t=i tik	P=D N P=D .	oui	non	non	non	non	non	non	non	non	non	non	non	non	non	oui
11	Art.Def	lolipe	-	-	o lolipe kaj si portokali	D N Rel	non	non	non	non	non	non	non	non	non	pas de A	pas de A	non	non	non	non
12	Art.Def +	kultùlke	-	-	pash t=o trin kultùlke	P=D Num N	non	non	non	non	non	non	non	pas d'A	non	non	non	non	non	non	non
13	Art.Def	làpsi	couleur + co	lolo + z	o làpsi... o... lolo em o	D N... D A &	oui	non	non	non	non	non	non	non	non	non	non	non	non	non	oui
14	Art.Def	manaqýri	couleur	parno	o manyqýri o parno	D N D A	oui	non	non	non	non	non	non	non	non	non	non	non	non	non	oui
15	Art.Def	-	couleur	parno	pash t=o parno	P=D A	non	non	non	non	non	non	non	non	non	non	pas de N	non	non	non	non
16	Art.Def	-	couleur	ròza	i li o ròza	Foc D A	non	non	non	non	non	non	non	non	non	non	pas de N	non	non	non	non
17	Art.Def	-	couleur	ròza	pash t=i ròza	P=D A	non	non	non	non	non	non	non	non	non	non	pas de N	non	non	non	non
18	Art.Def	kàpsa	couleur	ròza	i kàpsa i ròza	D N D A	oui	non	non	non	non	non	non	non	non	non	non	non	non	non	oui
19	Art.Def	manaqýri	couleur	parno	pash t=o... pash t=o pa	P=D A... N	non	non	oui	non	non	non	non	non	non	non	non	non	non	oui	non
20	Art.Def	kàpsa	-	-	t=i kàpsa	P=D N	non	non	non	non	non	non	non	non	non	non	pas de A	non	non	non	non
21	Art.Def	manaqýri	couleur	ròza	em t=i... t=o manyqýri t	Foc P=D N F	oui	non	non	non	non	non	non	non	non	non	non	non	non	non	oui
22	Art.Def	çeni	-	-	i çeni pash late	D N	non	non	non	non	non	non	non	non	non	non	pas de A	non	non	non	non
23	Quant	-	-	-	pash so trinende	P Quant Num	non	non	non	non	non	non	non	pas d'A	non	non	non	non	non	non	non
24	Art.Def	kàpsa	couleur	jeshìli	i kàpsa i jeshìli	D N D A	oui	non	non	non	non	non	non	non	non	non	non	non	non	non	oui
25	Art.Def	làpsi	couleur	sarìa	o sarìa làpsi	D A N	non	non	oui	non	non	non	non	non	non	non	non	non	non	oui	non
26	Art.Def	làpsi	couleur	kalo	em o kalo làpsi	Foc D A N	non	non	oui	non	non	non	non	non	non	non	non	non	non	oui	non

Figure 92 : Codage de l'analyse linguistique pour chaque occurrence du corpus spontané

3.2 Variation linguistique – données

Tableau 125 : Les structures et déterminants du corpus : avec quels adjectifs les trouve-t-on ?

Structures ou DET introducteurs	Adjectifs en romani, traduction		
DEF N DEF A	*banđi*, tordue *bar/bari/barit*, grand *ble/blu/bluja*, bleu *čhindi/čhinli*, coupée *gri/grija*, gris *ješ/ješil/ješili/ješilka*, vert *kalo*, noir *loli/lolo*, rouge *mesmer*, moyen	*parni/parno*, blanc *pičhindi*, non coupée *portokal/portokáli /portokalí/portokalia*, orange *ros/rosa/roz/roza*, rose *saria*, jaune *tikine/tikini/tikne /tikni/tikno/cikni*, petit *uci*, longue	*uštime*, pointu *vaver/vaverja*, autre *verð/vərði*, jaune *višnja*, rouge cerise *xurdi*, toute petite *zelino*, verte *žol/žolto*, jaune
DEF N A	*banđi*, tordue *bari/baro*, grand *ble/blu/bluja*, bleu *gri*, gris *ješil/ješili/šili*, vert *kafe*, café *kali/kalo*, noir *levo*, gauche	*loli/lolo*, rouge *parni/parno*, blanc *plavo*, bleu *portokali/portokalia*, orange *ros/roz/roza/roze*, rose *saria*, jaune *tikine/tikini/tikino /tikni/tikno/cikni*, petit	*uštime*, pointu *vaver/avaver*, autre *verð*, jaune *zeleni/zéleno/zeléno /zelino/želino/žilini*, verte *žolto*, jaune *zlato*, doré *zurlo/zurno*, jaune
DEF A N	*bangi/banđi*, tordu *bare/bari/baro*, grand *ble/bluja* , bleu *čhindi/čhindo/čhinli*, coupé *ješil/ješili/ješilka*, vert *kali/kalo*, noir *lačho*, beau/droit *levo*, gauche *loli/lolo*, rouge	*maškaruni*, centrale *parni/parno*, blanc *plavo*, bleu *pomarandža*, orange *portokaali*, orange *ros/roza*, rose *saji*, droit *sarija*, jaune *srednjo*, central	*tikne/tikni/tikno/tiynise /cikni/cikno/cikne/*, petit *uštime*, pointu *vaver/vavera/aver*, autre *džungalo*, dégoûtant *xurdi/xurdo*, tout petit *zeleni/zélino/zelíno*, vert *žolto*, jaune

complexes	*banđi*, tordue *bari/baro*, grand *čhindi*, coupée *fundi*, de derrière *ješili*, vert *kalo*, noir	*loli/lolo*, rouge *mesmja*, central *parni/parno*, blanc *roza*, rose *saria*, jaune *špicasto*, pointu	*tikini/tikni/tikno/tiɣnise*, petit *vaver*, autre *vərð/verða*, jaune *xurdi/xurdo*, tout petit *zelino*, vert *žolto*, jaune
DEM	*banđi*, tordue *bari*, grande *blu*, bleu *čhindi*, coupée *ješili*, jaune *kalo*, noir *loli/lolo*, rouge	*mesme/mesmer*, moyen *parni/parno*, blanc *portokali*, orange *prasina*, grise [poussiéreuse] *roza*, rose *saria*, jaune	*tikini/tikino/tikne /tikno/cikne/cikni*, petit *vaver/aver*, autre *vərði*, jaune *zeleno*, vert *žolto*, jaune
POSS	*desno*, droit	*majt/majti*, gauche	
IND	*banđi*, tordue *bari/baro*, grand *blu/bluja*, bleu *čhinli*, coupée *danderdi*, mordue *gri*, gris *ješili*, jaune *kali/kalo*, noir	*loli/lolo*, rouge *nevo*, neuf *parno*, blanc *portokáli/portokalí /portokalia*, orange *ros/roza*, rose *saria*, jaune *tikni/tikno/cikno*, petit	*uči*, longue *vaver/aver*, autre *xatardi*, pourrie *xurdi*, toute petite *zeleni/zelino*, vert *žolto*, jaune *zlato*, doré
NUM / QUANT	*aver*, autre *bari*, grande	*lolo*, rouge *tikni*, petite	*zeleno*, vert *žolto*, jaune
SGEN	*bango*, tordu *bar/bari/baro*, grand *ješil*, vert	*loli/lolo*, rouge *parni/parno*, blanc *roz*, rose	*tikni*, petite *vaver*, autre *zeleni*, verte
SREL	*bari/baro*, grand *čhindi/čhindo*, coupé *ješili*, vert *kalo*, noir *mašk...*, central	*parni/parno*, blanc *pičhindi*, non coupée *portokalia*, orange *roz/roza*, rose *tikini/tikino/tikni/tikno/cikni*, petit	*vaver/aver*, autre *xatardi*, pourrie *xurdi*, toute petite *zeleno/zelino*, vert *zlato*, doré

A seul (pas de N)	*bangi/banđi*, tordue *bare/bari/baro*, grand *ble/blu/bluja*, bleu *čhindi/čhinli/čhinlo*, coupé *crveno*, rouge *dijagonal*, diagonal *dur*, loin *fundi/fundit*, de derrière *gri/grija*, gris *ješili/ješilka/ješilt/šili*, vert *kalo*, noir *lačhi*, belle/droite	*levo*, gauche *loli/lolo*, rouge *majti*, gauche *mašk.../maškaruni /maškarutni*, centrale *parni/parno*, blanc *plavo*, bleu *portokalevo/portokáli /portokalí/portokalia*, orange *ros/roz/roza*, rose *saria*, jaune *škurte*, court	*špicasto*, pointu *tikini/tikino /tikni/cikni*, petit *uči*, longue *uštime*, pointu *vaver/vavera/aver*, autre *verð/verðe/verði*, jaune *vertikal*, vertical *xurdi*, toute petite *zeleno/zélino/zelíno/želini*, vert *žolto*, jaune *zlato*, doré
N seul (pas de A)	*aver*, autre *fundi*, de derrière *ješili*, vert	*kali*, noire *parno*, blanc *portokalia*, orange	*tikini/tikni*, petite *verði*, jaune
A N seul (sans déterminant)	*bari/baro*, grand *boši*, vide/blanc *gri*, gris *ješili*, vert *kalo*, noir *lačho*, beau/droit	*lolo*, rouge *marandža*, orange *parni/parno*, blanc *rozeva*, rose *tiɣnise/tikino/tikni*, petit	*vaver/vavera*, autre *džungalo*, dégoûtant *zelena/zeleno/zelino*, vert *zlato*, doré *žolto*, jaune

Tableau 126 : Les adjectifs du corpus : avec quelles structures ou déterminants les trouve-t-on ?

Adjectifs **Structures ou déterminants**	*bangi /bango/banđi*, tordu DEF N DEF A DEM DEF N A IND DEF A N SGEN complexes A seul	*bar/bare/bari/barit/baro*, grand DEF N DEF A NUM / QUANT DEF N A SGEN DEF A N SREL complexes A seul DEM A N seul IND	*ble/blu/bluja*, bleu DEF N DEF A DEM DEF N A IND DEF A N A seul	*čhindi/čhindo/čhinli/čhinlo*, coupé DEF N DEF A IND DEF A N SREL complexes A seul DEM
Adjectifs **Structures ou déterminants**	*fundi*, de derrière complexes N seul A seul	*gri/grija*, gris DEF N DEF A A seul DEF N A A N seul IND	*ješ/ješil/ješili/ješilka/ješilt/šili*, vert DEF N DEF A IND DEF N A SGEN DEF A N SREL complexes A seul DEM A N seul	*kali/kalo*, noir DEF N DEF A IND DEF A N SREL DEF A N A seul complexes N seul DEM A N seul
Adjectifs **Structures ou déterminants**	*lačhi/lačho*, beau/droit DEF A N A N seul A seul	*levo*, gauche DEF N A A seul DEF A N	*loli/lolo*, rouge DEF N DEF A IND A DEF N A NUM / QUANT DEF A N SGEN complexes A seul DEM A N seul	*mašk.../maškaruni/ maškarutni*, centrale DEF A N A seul SREL
Adjectifs **Structures ou déterminants**	*mesme/mesmer/mesmja*, moyen DEF N DEF A DEM complexes	*parni/parno*, blanc DEF N DEF A SGEN DEF N A SREL DEF A N A seul complexes N seul DEM A N seul IND	*plavo*, bleu DEF N A A seul DEF A N	*portokal/portokáli/portokalí/ portokalia/portokalevo*, orange DEF N DEF A IND DEF N A SREL DEF A N A seul DEM N seul
Adjectifs **Structures ou déterminants**	ros/roz/rosa/roza/ roze/rozeva, rose DEF N DEF A IND DEF N A SGEN DEF A N SREL complexes A seul DEM A N seul	saria, jaune DEF N DEF A DEM DEF N A IND DEF A N A seul complexes	tikine/tikini/tikino/tikne/tikni/ tikno/tiɣnise/cikne/cikni/cikno, petit DEF N DEF A NUM / QUANT DEF N A SGEN DEF A N SREL complexes A seul DEM N seul IND A N seul	uči, longue DEF N DEF A A seul IND
Adjectifs	*uštime*, pointu	*vaver/vavera/vaverja/ aver*, autre	*verð/verða/verðe/ verði/vərð/vərði*, jaune	*xurdi/xurdo*, tout petit

Structures ou déterminants	DEF N DEF A DEF N A DEF A N A seul	DEF N DEF A DEF N A DEF A N complexes DEM IND SGEN SREL A seul N seul A N seul	DEF N DEF A DEF N A complexes DEM A seul N seul	DEF N DEF A DEF A N complexes IND SREL A seul
Adjectifs **Structures ou déterminants**	*zelena/zeleni/zéleno/zeléno/zélino/zelíno/želino/želini/žilini*, vert DEF N DEF A DEF N A DEF A N complexes DEM IND NUM / QUANT SGEN SREL A seul A N seul	*zlato*, doré DEF N A IND SREL A seul AN seul	*žol/žolto*, jaune DEF N DEF A DEF N A DEF A N complexes DEM IND NUM / QUANT A seul A N seul	

Adjectifs **Structures ou déterminants**	*boši*, vide/blanc A N seul	*crveno*, rouge A seul	*danderdi*, mordue IND	*desno*, droit POSS	*dijagonal*, diagonal A seul	*dur*, loin A seul	*džungalo*, dégoûtant DEF A N A N seul
Adjectifs **Structures ou déterminants**	*kafe*, café DEF N A	*majt/majti*, gauche POSS A seul	*nevo*, neuf IND	*pomarandža/marandža*, orange DEF A N A N seul	*pičhindi*, non coupée DEF N DEF A SREL	*prasina*, grise DEM	*saji*, droit DEF A N
Adjectifs **Structures ou déterminants**	*srednjo*, central DEF A N	*škurte*, court A seul	*špicasto*, pointu complexes A seul	*vertikal*, vertical A seul	*višnja*, rouge cerise DEF N DEF A	*xatardi*, pourrie IND SREL	*zurlo/zurno*, jaune DEF N A

Tableau 127 : Structure des SN définis comprenant un adjectif produit par les participants des tests 1 et 2

	DEF N DEF A	DEF N A	DEF A N	complexes	Total
Test 1	81 (47,65 %)	49 (28,82 %)	35 (20,59 %)	5 (2,94 %)	170 (100 %)
Test 2	339 (35,65 %)	266 (27,97 %)	301 (31,65 %)	45 (4,73 %)	951 (100 %)
Total	420 (37,47 %)	315 (28,10 %)	336 (29,97 %)	50 (4,46 %)	1121 (100 %)

Tableau 128 : Structure des SN définis comprenant un adjectif produit par les témoins

	DEF N DEF A	DEF N A	DEF A N	complexes	Total
F24K	0 (0 %)	2 (11,76 %)	15 (88,24 %)	0 (0 %)	17 (100 %)
M27Dž	1 (5,26 %)	3 (15,79 %)	15 (78,95 %)	0 (0 %)	19 (100 %)
M43X	0 (0 %)	0 (0 %)	20 (86,96 %)	3 (13,04 %)	23 (100 %)
Total	1 (1,70 %)	5 (8,47 %)	50 (84,75 %)	3 (5,08 %)	59 (100 %)

Tableau 129 : Structure des SN définis et indéfinis comprenant un adjectif produit par les participants des tests 1 et 2

	IND N IND A	IND N A	IND A N	complexes	Total
Test 1	0 (0 %)	15 (75 %)	2 (10 %)	3 (15 %)	20 (100 %)
Test 2	0 (0 %)	21 (53,85 %)	16 (41,02 %)	2 (5,13 %)	39 (100 %)
Total	0 (0 %)	36 (61,02 %)	18 (30,51 %)	5 (8,47 %)	59 (100 %)

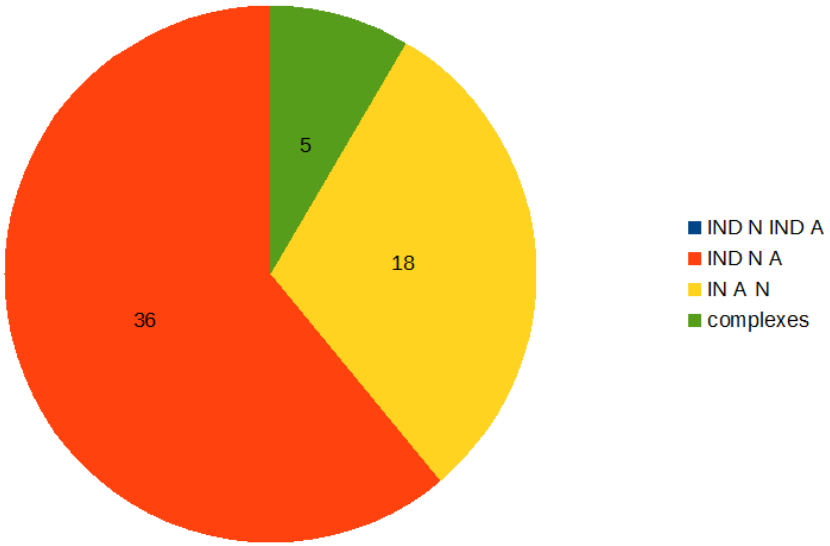


Figure 93 : Structure des SN définis et indéfinis comprenant un adjectif produit par les participants des tests 1 et 2

3.3 Variation sociale – tableaux

Tableau 130 : Structure des SN définis produits par les participants des tests 1 et 2 en fonction du GENRE

	DEF N DEF A	DEF N A	DEF A N	complexes	Total
Femmes	299 (48,38 %)	104 (16,83 %)	179 (28,96 %)	36 (5,83 %)	618 (100 %)
Hommes	121 (24,06 %)	211 (41,95 %)	157 (31,21 %)	14 (2,78 %)	503 (100 %)
Total	420 (37,47 %)	315 (28,10 %)	336 (29,97 %)	50 (4,46 %)	1121 (100 %)

Tableau 131 : Proportion d'adjectifs anté- et postposés produits par les participants des tests 1 et 2 en fonction du GENRE

	A postposés	A antéposés	Total
Femmes	464 (69,57 %)	203 (30,43 %)	667 (100 %)
Hommes	360 (67,77 %)	171 (32,20 %)	531 (100 %)
Total	824 (68,78 %)	374 (31,22 %)	1198 (100 %)

Tableau 132 : Structure des SN définis produits par les témoins en fonction du GENRE

	DEF N DEF A	DEF N A	DEF A N	complexes	Total
Femme	0 (0 %)	2 (11,76 %)	15 (88,24 %)	0 (0 %)	17 (100 %)
Hommes	1 (2,38 %)	3 (7,14 %)	35 (83,34 %)	3 (7,14 %)	42 (100 %)
Total	1 (1,70 %)	5 (8,47 %)	50 (84,75 %)	3 (5,08 %)	59 (100 %)

Tableau 133 : Proportion d'adjectifs anté- et postposés produits par les témoins en fonction du GENRE

	A postposés	A antéposés	Total
Femme	3 (11,54 %)	23 (88,46 %)	26 (100 %)
Hommes	7 (14,58 %)	41 (85,42 %)	48 (100 %)
Total	10 (13,51 %)	64 (86,49 %)	74 (100 %)

Tableau 134 : Structure des SN définis produits par les participants des tests 1 et 2 en fonction de la CLASSE D'AGE

	DEF N DEF A	DEF N A	DEF A N	complexes	Total
- 15 ans	16 (10,46 %)	128 (83,66 %)	9 (5,88 %)	0 (0 %)	153 (100 %)
15-39 ans	265 (50,09 %)	117 (22,12 %)	119 (22,50 %)	28 (5,29 %)	529 (100 %)
40-59 ans	115 (42,43 %)	62 (22,88 %)	83 (30,63 %)	11 (4,06 %)	271 (100 %)
+ 60 ans	24 (14,29 %)	8 (4,76 %)	125 (74,40 %)	11 (6,55 %)	168 (100 %)
Total	420 (37,47 %)	315 (28,10 %)	336 (29,97 %)	50 (4,46 %)	1121 (100 %)

Tableau 135 : Proportion d'adjectifs anté- et postposés produits par les participants des tests 1 et 2 en fonction de la CLASSE D'AGE

	A postposés	A antéposés	Total
- 15 ans	152 (94,41 %)	9 (5,59 %)	161 (100 %)
15-39 ans	421 (75,45 %)	137 (24,55 %)	558 (100 %)
40-59 ans	210 (70 %)	90 (30 %)	300 (100 %)
+ 60 ans	41 (22,91 %)	138 (77,09 %)	179 (100 %)
Total	824 (68,78 %)	374 (31,22 %)	1198 (100 %)

Tableau 136 : Structure des SN définis produits par les témoins en fonction de la CLASSE D'AGE

	DEF N DEF A	DEF N A	DEF A N	complexes	Total
15 – 39 ans	1 (2,78 %)	5 (13,89 %)	30 (83,33 %)	0 (0 %)	36 (100 %)
40 – 59 ans	0 (0 %)	0 (0 %)	20 (86,96 %)	3 (13,04 %)	23 (100 %)
Total	1 (1,70 %)	5 (8,47 %)	50 (84,75 %)	3 (5,08 %)	59 (100 %)

Tableau 137 : Proportion d'adjectifs anté- et postposés produits par les témoins en fonction de la CLASSE D'AGE

	A postposés	A antéposés	Total
15 – 39 ans	7 (14,29 %)	42 (85,71 %)	49 (100 %)
40 – 59 ans	3 (12 %)	22 (88 %)	25 (100 %)
Total	10 (13,51 %)	64 (86,49 %)	74 (100 %)

Tableau 138 : Structure des SN définis produits par les participants des tests 1 et 2 en fonction de l'EDUCATION

	DEF N DEF A	DEF N A	DEF A N	complexes	Total
0 à 5 années	106 (31,64 %)	35 (10,45 %)	172 (51,34 %)	22 (6,57 %)	335 (100 %)
6 à 12 années	213 (38,45 %)	212 (38,26 %)	114 (20,58 %)	15 (2,71 %)	554 (100 %)
+ 12 années	101 (43,54 %)	68 (29,31 %)	50 (21,55 %)	13 (5,60 %)	232 (100 %)
Total	420 (37,47 %)	315 (28,10 %)	336 (29,97 %)	50 (4,46 %)	1121 (100 %)

Tableau 139 : Proportion d'adjectifs anté- et postposés produits par les participants des tests 1 et 2 en fonction de l'EDUCATION

	A postposés	A antéposés	Total
0 à 5 années	171 (47,63 %)	188 (52,37 %)	359 (100 %)
6 à 12 années	464 (77,98 %)	131 (22,02 %)	595 (100 %)
+ 12 années	189 (77,46 %)	55 (22,54 %)	244 (100 %)
Total	824 (68,78 %)	374 (31,22 %)	1198 (100 %)

Tableau 140 : Structure des SN définis produits par les témoins en fonction de l'EDUCATION

	DEF N DEF A	DEF N A	DEF A N	complexes	Total
6 à 12 années	0 (0 %)	0 (0 %)	20 (86,96 %)	3 (13,04 %)	23 (100 %)
+ 12 années	1 (2,78 %)	5 (13,89 %)	30 (83,33 %)	0 (0 %)	36 (100 %)
Total	1 (1,70 %)	5 (8,47 %)	50 (84,75 %)	3 (5,08 %)	59 (100 %)

Tableau 141 : Proportion d'adjectifs anté- et postposés produits par les témoins en fonction de l'EDUCATION

	A postposés	A antéposés	Total
6 à 12 années	3 (12 %)	22 (88 %)	25 (100 %)
+ 12 années	7 (14,29 %)	42 (85,71 %)	49 (100 %)
Total	10 (13,51 %)	64 (86,49 %)	74 (100 %)

Tableau 142 : Structure des SN définis produits par les participants des tests 1 et 2 en fonction du MILITANTISME

	DEF N DEF A	DEF N A	DEF A N	complexes	Total
Militants	80 (30,08 %)	68 (25,56 %)	105 (39,47 %)	13 (4,89 %)	266 (100 %)
Non-militants	340 (39,76 %)	247 (29,89 %)	231 (27,02 %)	37 (4,33 %)	855 (100 %)
Total	420 (37,47 %)	315 (28,10 %)	336 (29,97 %)	50 (4,46 %)	1121 (100 %)

Tableau 143 : Proportion d'adjectifs anté- et postposés produits par les participants des tests 1 et 2 en fonction du MILITANTISME

	A postposés	A antéposés	Total
Militants	168 (60,65 %)	109 (39,35 %)	277 (100 %)
Non-militants	656 (71,23 %)	265 (28,77 %)	921 (100 %)
Total	824 (68,78 %)	374 (31,22 %)	1198 (100 %)

Tableau 144 : Structure des SN définis produits par les témoins en fonction du MILITANTISME

	DEF N DEF A	DEF N A	DEF A N	complexes	Total
Non-militant	1 (1,70 %)	5 (8,47 %)	50 (84,75 %)	3 (5,08 %)	59 (100 %)

Tableau 145 : Proportion d'adjectifs anté- et postposés produits par les témoins en fonction du MILITANTISME

	A postposés	A antéposés	Total
Non-militant	10 (13,51 %)	64 (86,49 %)	74 (100 %)

Tableau 146 : Structure des SN définis produits par les participants des tests 1 et 2 en fonction de la VARIETE

	DEF N DEF A	DEF N A	DEF A N	complexes	Total
Néo-locuteur arli	1 (2,27 %)	5 (11,37 %)	34 (77,27 %)	4 (9,09 %)	44 (100 %)
Mečkar	124 (62 %)	54 (27 %)	7 (3,50 %)	15 (7,50 %)	200 (100 %)
Arli	109 (29,46 %)	50 (13,51 %)	196 (52,97 %)	15 (4,06 %)	370 (100 %)
Čergar I	25 (50 %)	21 (42 %)	3 (6 %)	1 (2 %)	50 (100 %)
Arli & čergar I	141 (36,62 %)	156 (40,52 %)	75 (19,48 %)	13 (3,38 %)	385 (100 %)
Arli & bamlija	8 (32 %)	5 (20 %)	11 (44 %)	1 (4 %)	25 (100 %)

Mečkar & čergar I	4 (15,39 %)	12 (46,15 %)	10 (38,46 %)	0 (0 %)	26 (100 %)
Mečkar & bamlija	8 (38,10 %)	12 (57,14 %)	0 (0 %)	1 (4,76 %)	21 (100 %)
Total	420 (37,47 %)	315 (28,10 %)	336 (29,97 %)	50 (4,46 %)	1121 (100 %)

Tableau 147 : Proportion d'adjectifs anté- et postposés produits par les participants des tests 1 et 2 en fonction de la VARIETE

	A postposés	**A antéposés**	**Total**
Néo-locuteur arli	8 (18,18 %)	36 (81,82 %)	44 (100 %)
Mečkar	198 (95,19 %)	10 (4,81 %)	208 (100 %)
Arli	183 (46,45 %)	211 (53,55 %)	394 (100 %)
Čergar I	49 (94,23 %)	3 (5,77 %)	52 (100 %)
Arli & čergar I	331 (78,25 %)	92 (21,75 %)	423 (100 %)
Arli & bamlija	13 (54,17 %)	11 (45,83 %)	24 (100 %)
Mečkar & čergar I	20 (64,52 %)	11 (35,48 %)	31 (100 %)
Mečkar & bamlija	22 (100 %)	0 (0 %)	22 (100 %)
Total	824 (68,78 %)	374 (31,22 %)	1198 (100 %)

Tableau 148 : Structure des SN définis produits par les témoins en fonction de la VARIETE

	DEF N DEF A	**DEF N A**	**DEF A N**	**complexes**	**Total**
Kovački	0 (0 %)	2 (11,76 %)	15 (88,24 %)	0 (0 %)	17 (100 %)
Džambazo	1 (5,26 %)	3 (15,79 %)	15 (78,95 %)	0 (0 %)	19 (100 %)
Xoraxane	0 (0 %)	0 (0 %)	20 (86,96 %)	3 (13,04 %)	23 (100 %)
Total	1 (1,70 %)	5 (8,47 %)	50 (84,75 %)	3 (5,08 %)	59 (100 %)

Tableau 149 : Proportion d'adjectifs anté- et postposés produits par les témoins en fonction de la VARIETE

	A postposés	**A antéposés**	**Total**
Kovački	3 (11,54 %)	23 (88,46 %)	26 (100 %)
Džambazo	4 (17,39 %)	19 (82,61 %)	23 (100 %)
Xoraxane	3 (12 %)	22 (88 %)	25 (100 %)
Total	10 (13,51 %)	64 (86,49 %)	74 (100 %)

Tableau 150 : Regroupement des modalités convergentes issues de GENRE, CLASSE D'AGE, EDUCATION et VARIETE dans le pourcentage de production des différents SN définis

	DEF N DEF A	DEF N A	DEF A N	complexes	Total
Témoins	1 (1,70 %)	5 (8,47 %)	50 (84,75 %)	3 (5,08 %)	59 (100 %)
- 15 ans	16 (10,46 %)	128 (83,66 %)	9 (5,88 %)	0 (0 %)	153 (100 %)
+ 60 ans	24 (14,29 %)	8 (4,76 %)	125 (74,40 %)	11 (6,55 %)	168 (100 %)
Arlis	109 (29,46 %)	50 (13,51 %)	196 (52,97 %)	15 (4,06 %)	370 (100 %)
0 à 5 années	106 (31,64 %)	35 (10,45 %)	172 (51,34 %)	22 (6,57 %)	335 (100 %)
Arlis & čergars I	141 (36,62 %)	156 (40,52 %)	75 (19,48 %)	13 (3,38 %)	385 (100 %)
40-59 ans	115 (42,43 %)	62 (22,88 %)	83 (30,63 %)	11 (4,06 %)	271 (100 %)
Femmes	299 (48,38 %)	104 (16,83 %)	179 (28,96 %)	36 (5,83 %)	618 (100 %)
Čergars I	25 (50 %)	21 (42 %)	3 (6 %)	1 (2 %)	50 (100 %)
15-39 ans	265 (50,09 %)	117 (22,12 %)	119 (22,50 %)	28 (5,29 %)	529 (100 %)
Mečkars	124 (62 %)	54 (27 %)	7 (3,50 %)	15 (7,50 %)	200 (100 %)

Tableau 151 : Regroupement des modalités convergentes issues de GENRE, CLASSE D'AGE, EDUCATION et VARIETE dans le pourcentage de production d'adjectifs anté- et postposés

	A postposés	A antéposés	Total
Témoins	10 (13,51 %)	64 (86,49 %)	74 (100 %)
+ 60 ans	41 (22,91 %)	138 (77,09 %)	179 (100 %)
Arlis	183 (46,45 %)	211 (53,55 %)	394 (100 %)
0 à 5 années	171 (47,63 %)	188 (52,37 %)	359 (100 %)
Femmes	464 (69,57 %)	203 (30,43 %)	667 (100 %)
40-59 ans	210 (70 %)	90 (30 %)	300 (100 %)
15-39 ans	421 (75,45 %)	137 (24,55 %)	558 (100 %)
Arlis & čergars I	331 (78,25 %)	92 (21,75 %)	423 (100 %)
Čergars I	49 (94,23 %)	3 (5,77 %)	52 (100 %)
- 15 ans	152 (94,41 %)	9 (5,59 %)	161 (100 %)
Mečkars	198 (95,19 %)	10 (4,81 %)	208 (100 %)

Tableau 152 : Pourcentage de production des différents SN définis par les participants des tests 1 et 2 en fonction de INDIVIDU

	DEF N DEF A	DEF N A	DEF A N	complexes	Total
M43X	0 (0 %)	0 (0 %)	20 (86,96 %)	3 (13,04 %)	23 (100 %)
M71A	0 (0 %)	0 (0 %)	31 (100 %)	0 (0 %)	31 (100 %)
F24K	0 (0 %)	2 (11,76 %)	15 (88,24 %)	0 (0 %)	17 (100 %)
M22a	1 (2,27 %)	5 (11,36 %)	34 (77,27 %)	4 (9,09 %)	44 (100 %)
F65A	2 (2,60 %)	1 (1,30 %)	70 (90,91 %)	4 (5,19 %)	77 (100 %)
M27Dž	1 (5,26 %)	3 (15,79 %)	15 (78,95 %)	0 (0 %)	19 (100 %)
M11A	1 (6,25 %)	15 (93,75 %)	0 (0 %)	0 (0 %)	16 (100 %)
M19M$_{(1)}$	1 (6,67 %)	13 (86,67 %)	1 (6,67 %)	0 (0 %)	15 (100 %)
M13AČ	13 (9,70 %)	112 (83,58 %)	9 (6,72 %)	0 (0 %)	134 (100 %)
M58Č	3 (13,64 %)	19 (86,36 %)	0 (0 %)	0 (0 %)	22 (100 %)
M46A	10 (14,29 %)	13 (18,57 %)	46 (65,71 %)	1 (1,43 %)	70 (100 %)
M58MČ	4 (14,38 %)	12 (46,15 %)	10 (38,46 %)	0 (0 %)	26 (100 %)
M34AČ	3 (25 %)	0 (0 %)	8 (66,67 %)	1 (8,33 %)	12 (100 %)
F85A	9 (30 %)	0 (0 %)	17 (56,67 %)	4 (13,33 %)	30 (100 %)
F24AČ	7 (30,44 %)	4 (17,39 %)	12 (52,17 %)	0 (0 %)	23 (100 %)
M24	8 (32 %)	5 (20 %)	11 (44 %)	1 (4 %)	25 (100 %)
F60A	4 (36,36 %)	0 (0 %)	6 (54,55 %)	1 (9,09 %)	11 (100 %)
F27A	6 (37,5 %)	3 (18,75 %)	7 (43,75 %)	0 (0 %)	16 (100 %)
F37MB	8 (38,10 %)	12 (57,14 %)	0 (0 %)	1 (4,76 %)	21 (100 %)
M37A	2 (40 %)	1 (20 %)	0 (0 %)	2 (40 %)	5 (100 %)
F38A	16 (42,11 %)	7 (18,42 %)	15 (39,47 %)	0 (0 %)	38 (100 %)
F66M	7 (46,67 %)	6 (40 %)	0 (0 %)	2 (13,33 %)	15 (100 %)
M60A	2 (50 %)	1 (25 %)	1 (25 %)	0 (0 %)	4 (100 %)
F11A	1 (50 %)	1 (50 %)	0 (0 %)	0 (0 %)	2 (100 %)
F44AČ	35 (52,24 %)	4 (5,97 %)	23 (34,33 %)	5 (7,46 %)	67 (100 %)
F22AČ	7 (53,85 %)	4 (30,77 %)	2 (15,38 %)	0 (0 %)	13 (100 %)
F16AČ	64 (55,65 %)	31 (26,96 %)	13 (11,30 %)	7 (6,09 %)	115 (100 %)
F23AČ	12 (57,14 %)	1 (4,76 %)	8 (38,09 %)	0 (0 %)	21 (100 %)
M52A	6 (60 %)	2 (20 %)	2 (20 %)	0 (0 %)	10 (100 %)
M19M$_{(2)}$	23 (63,89 %)	8 (22,22 %)	1 (2,78 %)	4 (11,11 %)	36 (100 %)
F22M	42 (64,62 %)	15 (23,08 %)	4 (6,15 %)	4 (6,15 %)	75 (100 %)
F47M	29 (65,91 %)	9 (20,46 %)	1 (2,27 %)	5 (11,36 %)	44 (100 %)
M18Č	22 (78,57 %)	2 (7,14 %)	3 (10,71 %)	1 (3,57 %)	28 (100 %)
F20A	43 (82,69 %)	6 (11,54 %)	0 (0 %)	3 (5,77 %)	52 (100 %)
F53A	6 (85,71 %)	0 (0 %)	1 (14,29 %)	0 (0 %)	7 (100 %)
M45M	22 (88 %)	3 (12 %)	0 (0 %)	0 (0 %)	25 (100 %)
F9A	1 (100 %)	0 (0 %)	0 (0 %)	0 (0 %)	1 (100 %)
Total	421 (35,68 %)	319 (27,03 %)	387 (32,80 %)	53 (4,49 %)	1180 (100 %)

Tableau 153 : Pourcentage de production d'adjectifs anté- et postposés par les participants des tests 1 et 2 en fonction de INDIVIDU

	A postposés	A antéposés	Total
M71A	0 (0 %)	37 (100 %)	37 (100 %)
F65A	5 (6,49 %)	72 (93,51 %)	77 (100 %)
F24K	3 (11,54 %)	23 (88,46 %)	26 (100 %)
M43X	3 (12 %)	22 (88 %)	25 (100 %)
M27Dž	4 (17,39 %)	19 (82,61 %)	23 (100 %)
M22a	8 (18,18 %)	36 (81,82 %)	44 (100 %)
M34AČ	3 (27,27 %)	8 (72,73 %)	11 (100 %)
F60A	4 (36,36 %)	7 (63,64 %)	11 (100 %)
F85A	11 (36,67 %)	19 (63,33 %)	30 (100 %)
M46A	28 (37,33 %)	47 (62,67 %)	75 (100 %)
F24AČ	21 (48,84 %)	22 (51,16 %)	43 (100 %)
M24	13 (54,17 %)	11 (45,83 %)	24 (100 %)
F27A	9 (56,25 %)	7 (43,75 %)	16 (100 %)
F38A	22 (57,89 %)	16 (42,11 %)	38 (100 %)
F23AČ	14 (63,64 %)	8 (36,36 %)	22 (100 %)
M58MČ	20 (64,52 %)	11 (35,48 %)	31 (100 %)
F44AČ	46 (64,79 %)	25 (35,21 %)	71 (100 %)
M60A	7 (70 %)	3 (30 %)	10 (100 %)
F22AČ	12 (80 %)	3 (20 %)	15 (100 %)
M52A	9 (81,82 %)	2 (18,18 %)	11 (100 %)
F16AČ	109 (86,51 %)	17 (13,49 %)	126 (100 %)
F53A	7 (87,50 %)	1 (12,50 %)	8 (100 %)
M18Č	25 (89,29 %)	3 (10,71 %)	28 (100 %)
M19M$_{(1)}$	14 (93,33 %)	1 (6,67 %)	15 (100 %)
M13AČ	126 (93,33 %)	9 (6,67 %)	135 (100 %)
M45M	31 (93,94 %)	2 (6,06 %)	33 (100 %)
F22M	63 (94,03 %)	4 (5,97 %)	67 (100 %)
F47M	45 (95,74 %)	2 (4,26 %)	47 (100 %)
M19M$_{(2)}$	31 (96,875 %)	1 (3,125 %)	32 (100 %)
M37A	5 (100 %)	0 (0 %)	5 (100 %)
M11A	16 (100 %)	0 (0 %)	16 (100 %)
F20A	50 (100 %)	0 (0 %)	50 (100 %)
F11A	4 (100 %)	0 (0 %)	4 (100 %)
F37MB	22 (100 %)	0 (0 %)	22 (100 %)
F66M	14 (100 %)	0 (0 %)	14 (100 %)
F9A	6 (100 %)	0 (0 %)	6 (100 %)
M58Č	24 (100 %)	0 (0 %)	24 (100 %)
Total	833 (65,49 %)	439 (34,51 %)	1272 (100 %)

4. Annexes du chapitre VI : Discussion

Tableau 154 : Profil des locuteurs ayant prononcé les exemples spontanés du chapitre VI

Nom	Age	Genre	Lieu d'enreg.	Lieu de naissance / de résidence[1]	Variété	Autres langues connues[2]	Nombre années scolarité	Profession	Militantisme
M27M	27	M	Llakatund (Albanie)	Berat / Llakatund (Albanie)	mečkar	albanais (+ italien, grec)	0 à 5 années	vendeur de fleurs	non
M30M	30	M	Voskopojë (Albanie)	Fier (Albanie)	mečkar	albanais, anglais (+ italien, grec)	plus 12 années	étudiant, travailleur social, militant, pianiste, guitariste	oui
M50A	50	M	Voskopojë (Albanie)	Tiranë (Albanie)	arli	albanais	6 à 12 années	travailleur social, militant	oui
M63GSb	63	M	Zemun (Serbie)	Leskovac / Belgrade, Zemun (Serbie)	gurbet	serbe, macédonien, bulgare, russe (+ anglais, allemand, hongrois)	plus 12 années	violoniste, enseignant de romani, poète	oui
M20VlBg	20	M	Thessalonique (Grèce)	Teteven (Bulgarie)	? (*vlax*)	bulgare, grec	6 à 12 années	sans emploi	non
M41VlGr	41	M	Larissa (Grèce)	Larissa (Grèce)	? (*vlax*)	grec, beaş-rudar (+ anglais)	0 à 5 années	travailleur social, militant	oui

[1] Nous mentionnons derrière la barre oblique le lieu de résidence si celui-ci est ou a été différent du lieu de naissance.

[2] La mention des langues connues autres que le romani correspond aux pratiques déclarées des locuteurs. J'ai distingué deux niveaux de connaissance : plutôt fluide ou bien plutôt partielle (indiquée entre parenthèses).