



HAL
open science

Les expressions référentielles et les chaînes de référence en français et en coréen

Seo Yeoung Kim

► **To cite this version:**

Seo Yeoung Kim. Les expressions référentielles et les chaînes de référence en français et en coréen. Linguistique. Université de Strasbourg, 2023. Français. NNT : 2023STRAC009 . tel-04295051

HAL Id: tel-04295051

<https://theses.hal.science/tel-04295051v1>

Submitted on 20 Nov 2023

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.

ÉCOLE DOCTORALE des Humanités

UR 1339 LiLPa

THÈSE présentée par :

Seo Yeoung KIM

soutenue le : **22 septembre 2023**

pour obtenir le grade de : **Docteur de l'université de Strasbourg**

Discipline/ Spécialité : Sciences du langage - Linguistique

**Les expressions référentielles et les chaînes
de référence en français et en coréen**

THÈSE dirigée par :

Madame SCHNEDECKER Catherine
Monsieur SHIM Bong-Sup

Professeure, Université de Strasbourg, co-directrice
Professeur, Université Nationale de Séoul, co-directeur

RAPPORTEURS :

Madame FREYERMUTH Sylvie
Madame ROSSI Caroline

Professeure, Université du Luxembourg
Professeure, Université Grenoble Alpes

AUTRES MEMBRES DU JURY :

Madame CHOI-JONIN Injoo
Madame KUYUMCUYAN Annie
Monsieur DIAS Dominique

Professeure, Université Toulouse Jean Jaurès
Professeure, Université de Strasbourg
Maître de Conférences, Sorbonne Université

Aux lecteurs

Remerciements

Ma première pensée s'adresse à ma directrice de thèse, Madame Catherine Schnedecker, sans qui ce travail n'aurait pas pu voir le jour. Depuis notre première rencontre en 2017, elle m'a guidée et encouragée tout au long de ce chemin avec ses conseils judicieux et sa disponibilité constante. À ses côtés, j'ai appris la valeur de la patience et de la rigueur. Cela a été une grande fierté et un plaisir de travailler sous sa direction.

Je souhaite également exprimer ma gratitude à Monsieur Bong-Sup Shim, mon co-directeur, d'avoir co-encadré et de m'avoir toujours encouragée. Sa présence et ses commentaires ont rendu mon travail enrichissant.

J'adresse mes sincères remerciements aux membres de mon jury. Un grand merci à Madame Sylvie Freyermuth, Professeure à l'Université du Luxembourg et à Madame Caroline Rossi, Professeure à l'Université Grenoble-Alpes qui ont accepté de donner de leur temps pour lire ma thèse et d'être les rapporteurs. Je remercie aussi Mme Annie Kuyumcuyan, Professeure à l'Université de Strasbourg, Madame Choi-Jonin, Professeure à l'Université Toulouse Jean Jaurès et Monsieur Dominique Dias, Maître de Conférences à la Sorbonne Université, d'avoir accepté d'évaluer mon travail.

Mes remerciements vont tout naturellement à mes collègues de la Villa Brunner, Chang, Marine, Bruno, Seto, Delphine, Cindy, Emmanuelle et Ioana, qui m'ont accueillie chaleureusement et aidée de près et de loin tout au long de ces dernières années. Une des plus grandes chances et des plus grandes joies dans mon quotidien à Strasbourg était mes collègues.

Je remercie tout particulièrement Jihyang qui m'a témoigné son amitié tout au long de ce trajet. Sa présence à mes côtés a été toujours encourageante et précieuse. J'espère pouvoir continuer notre aventure en Corée.

Un grand merci va également à mes amies, Lara, Chrissy, Hyo-jung, Hyein, Youngnan et Jeongmi. Grâce à elles, ma vie en France était plus intéressante et enrichissante.

Je souhaite remercier mes collègues francophones de SNU, Cécile et Rémi qui m'ont, de bon cœur, aidée et encouragée lors des dernières étapes de la rédaction de cette thèse.

J'ai une pensée particulière pour ma famille, mon mari, mes parents, ma sœur, mes beaux-parents et ma belle-sœur qui m'ont encouragée et soutenue pour arriver au bout de ce long travail. Je pense aussi à ma tante qui nous a quittés récemment et à ma fille qui vient d'arriver au monde.

Table des matières

Remerciements	i
Table des matières	ii
Liste des tableaux.....	ix
Liste des figures.....	x
Règle de translittération	xi

Introduction	1
--------------------	---

Première partie : Description générale des chaînes de référence

Chapitre 1 Problème de la cohérence et quelques notions fondamentales

Introduction.....	6
1 Problème de la cohérence et de la cohésion.....	7
1.1 La cohérence et la cohésion.....	7
1.2 Quatre plans d'organisation textuelle	10
2 Approches de la référence	13
2.1 Importance et aspects du problème	14
2.2 La notion de <i>référence</i>	16
2.2.1 Référence et existence	16
2.2.2 Référence et sens	18
2.2.2.1 Le triangle sémiotique et la référence externe.....	18
2.2.2.2 Sens, dénotation et référence	19
2.2.2.3 La saillance des référents	21
2.3 Diverses approches linguistiques du problème de la référence.....	24
2.3.1 Les approches en sémantique grammaticale référentielle	25
2.3.2 Les approches discursivo-fonctionnelles	26
2.3.3 Les approches configurationnelles ou relationnelles	27
3 Anaphore, coréférence et chaînes de référence	28
3.1 Anaphore et coréférence	29
3.1.1 Autonomie référentielle	29
3.1.2 Relation symétrique des entités	30
3.2 Chaînes de référence	31
3.2.1. La notion de <i>chaîne de référence</i>	31
3.2.2. Les différences entre anaphore, coréférence et chaîne de référence	33

Bilan.....	34
------------	----

Chapitre 2 Caractéristiques des chaînes de référence

Introduction.....	36
1 Les caractéristiques des chaînes.....	37
1.1 Les chaînes de référence et leurs bornes.....	37
1.2 Les chaînes de référence et leur déroulement dans le texte.....	40
2 Les caractéristiques des maillons.....	43
2.1 Expressions référentielles susceptibles de servir de maillon.....	43
2.2 Diversité des types de maillons et caractéristiques des chaînes.....	45
3 Trois facteurs relatifs à la variété du matériau lexical.....	46
3.1 Le genre textuel.....	47
3.2 L'époque de rédaction.....	50
3.3 La langue de rédaction.....	52
4 Un ensemble de paramètres pour caractériser les chaînes de référence.....	57
4.1 Les paramètres et leur signification.....	57
4.2 La fonction des paramètres dans l'analyse.....	62
Bilan.....	64

Chapitre 3 Les chaînes de référence d'un point de vue contrastif

Introduction.....	66
1 Approche inter-langues : comment refléter les différences ?.....	67
1.1 Approche inter-langues dans la linguistique française.....	67
1.2 Approche inter-langues dans la linguistique coréenne.....	69
2 Les particularités du coréen à considérer dans le domaine de la référence.....	71
2.1 Description générale de la langue coréenne : une langue agglutinante.....	71
2.2 L'absence de l'article et l'existence d'un système indirect et allusif.....	73
2.2.1 Rôle et fonctionnalité discursive de la particule en coréen.....	75
2.2.2 Influence du prédicat : le sens et la forme de prédicat.....	81
2.2.2.1 Le sens de prédicat.....	81
2.2.2.2 La forme de prédicat : choix de terminaison.....	82
2.3 La dépendance contextuelle ou situationnelle.....	85
Bilan.....	86

Chapitre 4 Les expressions référentielles en français et en coréen

Introduction.....	88
1 Le classement des expressions référentielles.....	89
1.1 Les expressions référentielles et le type de référence.....	89

1.2 Les expressions référentielles dans le cadre de la chaîne de référence	90
2 Les noms propres	91
2.1 La signification de nom propre	92
2.2 Caractéristiques et fonctions référentielles du nom propre	95
2.3 Les différentes formes des noms propres et leurs emplois.....	98
2.3.1 Les noms propres non modifiés et l'emploi prototypique.....	98
2.3.2 Les noms propres modifiés et les emplois non-prototypiques.....	100
2.3.2.1 Type dénominatif	100
2.3.2.2 Type métaphorique.....	102
2.3.2.3 Type métonymique	104
2.3.2.4 Type dit « de fractionnement »	105
2.4 Les emplois des noms propres en coréen	107
2.4.1 La définition des <i>noms propres</i> en linguistique coréenne : l'emploi prototypique	107
2.4.2 Les emplois non-prototypiques	110
2.4.2.1 Type dénominatif et éléments grammaticaux.....	110
2.4.2.2 Type métaphorique et éléments grammaticaux	113
2.4.2.3 Type métonymique et éléments grammaticaux.....	115
2.4.2.4 Type dit de fractionnement et éléments grammaticaux.....	117
2.5 Bilan.....	119
3 Les pronoms personnels.....	120
3.1 Les formes morphologiques variées du pronom personnel	121
3.2 La fonction référentielle des pronoms personnels	122
3.2.1 Les pronoms de première et de seconde personnes.....	122
3.2.1.1 Le pronom <i>je</i>	123
3.2.1.2 Le pronom <i>tu</i>	125
3.2.1.3 Les pronoms <i>nous</i> et <i>vous</i>	126
3.2.2 Les pronoms de troisième personne	128
3.3 Les pronoms personnels en coréen.....	131
3.3.1 Les emplois des pronoms personnels et leurs synonymes	132
3.3.2 Pronoms invariables et leurs rôles syntaxiques.....	135
3.3.3 Pronoms réservés à l'écrit.....	135
3.3.4 La fréquence d'emploi des pronoms zéros	136
3.3.5 Pronoms personnels et épithète	138
3.4 Bilan.....	138
4 Les syntagmes nominaux	139
4.1 Les expressions nominales définies	140
4.1.1 Les emplois des SN définis.....	140
4.1.2 Les SN définis complets et incomplets	140

4.2 Les expressions nominales indéfinies.....	142
4.2.1 Les emplois des SN indéfinis.....	142
4.2.2 La procédure d'interprétation des SN indéfinis.....	143
4.3 Les expressions nominales démonstratives.....	144
4.3.1 Les emplois des SN démonstratifs.....	144
4.3.2 L'interprétation référentielles des SN démonstratifs.....	145
4.4 Les expressions nominales du coréen.....	147
4.4.1 <i>Gwanhyungsa</i> + N.....	148
4.4.2 La construction possessive + N.....	149
4.4.3 N + N.....	149
4.4.4 La construction relative + N.....	151
4.4.5 Adjectif + N.....	151
4.5 Bilan.....	152

Deuxième partie : Méthodologie et conventions d'annotation

Chapitre 5 Constitution du corpus et outils d'analyse

Introduction.....	154
1 Corpus : choix de textes.....	155
1.1 Choix du genre textuel et la constitution du corpus.....	155
1.2 Présentation des textes du corpus.....	156
1.2.1 Présentation du roman français : <i>Le fantôme de l'Opéra</i>	157
1.2.2 Présentation du roman coréen : <i>Arbre à feuillage persistant</i>	158
1.3 Aperçu sur la construction textuelle.....	159
1.3.1 Structure textuelle de roman.....	160
1.3.2 Schéma de la structure narrative du roman.....	160
1.3.3 La structure et les personnages du <i>Fantôme</i>	161
1.3.4 La structure et les personnages de l' <i>Arbre</i>	163
2 Travail d'annotation.....	165
2.1 La fonction et le processus de l'annotation.....	165
2.2 Choix d'un outil d'annotation.....	166
2.3 Annotation avec SACR.....	167
3 Analyse des données <i>via</i> CR Viewer.....	170
3.1 L'interface du logiciel.....	170
3.1.1 « Select options » : différents types des données.....	171
3.1.2 « Display options » : conditions d'affichage des données.....	172
3.2 Deux aspects d'analyse.....	172

3.2.1 Tendance globale des chaînes de référence	173
3.2.2 Examen du caractère individuel des chaînes de référence	174
Bilan	176

Chapitre 6 Les conventions d'annotation

Introduction.....	178
1 Base de démarche d'annotation	179
1.1 Délimiter une expression référentielle.....	179
1.2 Traitement des singletons	180
1.3 Le manuel DEMOCRAT et notre position.....	180
2 Annotation des expressions référentielles	181
2.1 Types de référents et de relation coréférentielle	182
2.2 Les noms et les SN	186
2.2.1 Les expressions figées.....	186
2.2.2 Les noms sans déterminant	187
2.2.2.1 Référents actualisés et non actualisés.....	188
2.2.2.2 Fonction classifiante.....	189
2.2.3 La distinction des référents génériques et spécifiques	190
2.2.4 Le dédoublement du collectif	191
2.2.5 Les syntagmes nominaux attributifs.....	192
2.2.6 Les appositions nominales.....	194
2.2.6.1 Constructions appositives intégrées	194
2.2.6.2 Constructions appositives détachées.....	194
2.2.7 Les anaphores résomptives	196
2.3 Les pronoms.....	197
2.3.1 Pronoms personnels déictiques employés en différents niveaux discursifs.....	198
2.3.2 Pronom réfléchi <i>se</i> du français.....	199
2.3.3 Pronom <i>il</i> impersonnel du français.....	200
2.3.4 Pronom <i>on</i> du français	200
2.3.4.1 <i>On</i> et ses diverses interprétations.....	200
2.3.4.2 La forme <i>l'on</i>	202
2.3.5 Pronoms relatifs.....	202
2.3.6 Pronoms démonstratifs et constructions pseudo-clivées	203
2.3.6.1 Forme simple et forme renforcée	204
2.3.6.2 <i>C'est</i> et les constructions pseudo-clivées	205
2.4 Les verbes	206
2.4.1 Coordination des verbes	206
2.4.2 Destinataire de l'expression impérative	207

2.5 Autres cas.....	208
2.5.1 Coordination et délimitation des groupes.....	208
2.5.2 Les éléments partageant le même contenu lexical.....	208
2.5.3 Titre du texte et titre de chaque chapitre.....	210
2.5.4 Les nombres.....	211
3 Schéma des propriétés	211
3.1 Critères commun	211
3.1.1 Critère catégoriel	212
3.1.2 Critère référentiel.....	214
3.2 Critères supplémentaires	214
3.3 Schémas réalisés	216
4 Résultats escomptés.....	217
Bilan	218

Troisième partie : Synthèses des analyses

Chapitre 7 Analyse des données

Introduction.....	221
1 Analyse quantitative générale	222
1.1 Aperçu sur les chiffres-clés.....	222
1.2 Corpus français	223
1.3 Corpus coréen	225
2 Les différences entre langues dans la composition des chaînes	226
2.1 L'interférence potentielle et le choix des expressions référentielles.....	226
2.2 Usages des pronoms dans les deux langues.....	230
2.2.1 Les contextes d'apparition du pronom zéro.....	231
2.2.2 Usage restreint du pronom personnel en coréen.....	233
2.3 Usages des noms propres dans les deux langues.....	235
2.3.1 Les contextes d'apparition du nom propre non modifié	235
2.3.2 Les différents types de nom propre modifié.....	240
2.4 Usages des SN anaphoriques dans les deux langues.....	243
2.4.1 Les contextes d'apparition du SN nu	244
2.4.2 Les différents types de SN anaphorique.....	247
2.5 Les points de fraction et les marqueurs référentiels.....	249
3 Les chaînes de référence et l'influence du genre textuel	251
3.1 L'influence du genre textuel : le cas du texte encyclopédique	252
3.1.1 Aperçu des caractéristiques du texte encyclopédique.....	252
3.1.2 Les chaînes de référence dans les textes encyclopédiques	254

3.1.2.1 L'emploi du pronom	254
3.1.2.2 L'emploi des marqueurs référentiels de basse accessibilité	257
3.2 État de l'art sur les chaînes de référence en contexte narratif.....	263
3.3 Caractéristiques typiques des chaînes de référence dans le texte narratif	266
3.3.1 Le nombre de singletons	266
3.3.2 La longueur des chaînes	267
3.4 Le contexte spatio-temporel et les chaînes de référence	269
3.5 Les chaînes de référence des personnages dans les romans	273
3.5.1 Modalités d'introduction des personnages.....	273
3.5.1.1 Les maillons initiaux des chaînes des personnages.....	274
3.5.1.2 Le niveau énonciatif d'apparition de personnages	278
3.5.2 Les anaphores nominales des chaînes des personnages.....	284
4 Synthèse	287
Conclusion	290
Bibliographie	294

Liste des tableaux

Tableau 1 : Extrait d'un récapitulatif des différences entre anaphore, coréférence et CR (Schnedecker, 2019)	34
Tableau 2 : Distribution des expressions anaphoriques (Tutin, 2002)	48
Tableau 3 : Récapitulatif de la distribution des ER (Baumer, 2015)	54
Tableau 4 : Synthèse des paramètres caractérisant les CR (Schnedecker, 2021 : 109)	63
Tableau 5 : Structure sous-jacente de l'énoncé selon le choix de la particule	79
Tableau 6 : Catégories et variations morphologiques (GMF, 2018 : 368)	121
Tableau 7 : Les pronoms personnels en coréen	132
Tableau 8 : Tableau récapitulatif du corpus	156
Tableau 9 : Quatre types de relation « near-identity » et leurs sous-types (Recasens <i>et al.</i> , 2010) .	185
Tableau 10 : Tableau récapitulatif des chiffres-clés des deux corpus	222
Tableau 11 : Les occurrences des catégories grammaticales (<i>Fantôme</i>)	224
Tableau 12 : Les occurrences des catégories grammaticales (<i>Arbre</i>)	225
Tableau 13 : La hiérarchie des composants des CR dans les deux langues	226
Tableau 14 : Répartition des pronoms utilisés dans le corpus	231
Tableau 15 : La comparaison des sous-catégories de nom propres dans les deux corpus	235
Tableau 16 : La catégorisation des SN anaphoriques selon le degré d'accessibilité	248
Tableau 17 : La composition des SN anaphoriques des deux textes	248
Tableau 18 : Tableau récapitulatif des corpus utilisés dans les recherches antérieures	264
Tableau 19 : Caractéristiques et contenus de notre corpus	265
Tableau 20 : Nombre de chaînes selon la longueur	268
Tableau 21 : Les CR de plus de 30 maillons (<i>Fantôme</i>)	268
Tableau 22 : Les CR de plus de 30 maillons (<i>Arbre</i>)	269
Tableau 23 : Les éléments susceptibles d'influencer la longueur des chaînes	271
Tableau 24 : Expressions référentielles et leur position possible dans la chaîne (Schnedecker, 2021)	274
Tableau 25 : Le premier et le deuxième maillon de chaque chaîne sur le personnage	275
Tableau 26 : Différenciation des premières mentions selon le niveau énonciatif d'apparition	279
Tableau 27 : Le coefficient de stabilité des chaînes des personnages principaux	285
Tableau 28 : La redénomination <i>via</i> le nom propre	286

Liste des figures

Figure 1 : La schématisation des plans d'organisation textuelle (Charolles, 1988 : 10) appliquée à l'exemple (2)	13
Figure 2 : Triangles de C. Ogden & I. Richards (1923) et de S. Ullmann (1952), d'après Kleiber (1997)	18
Figure 3 : Anaphore coréférentielle et anaphore associative (Webber, 1988)	28
Figure 4 : Echelle de l'accessibilité (Ariel, 1990)	44
Figure 5 : Processus de la référence des Npr (Wilmet, 1991)	97
Figure 6 : Echelle d'indexicalité (Charolles, 2002)	127
Figure 7 : Les emplois des SN indéfinis (Charolles, 2002)	143
Figure 8 : Schéma narratif du roman	161
Figure 9 : Le schéma narratif des trois premiers chapitres du <i>Fantôme</i>	162
Figure 10 : Les personnages dans le <i>Fantôme</i>	163
Figure 11 : Le schéma narratif des trois premiers chapitres de <i>l'Arbre</i>	164
Figure 12 : Les personnages dans <i>l'Arbre</i>	164
Figure 13 : SACR : Délimitation des expressions référentielles	168
Figure 14 : SACR : Regroupement des coréférences	168
Figure 15 : SACR : Attribution des propriétés	169
Figure 16 : SACR : Utilisation du « Search box »	169
Figure 17 : CR Viewer : l'interface principale	170
Figure 18 : CR Viewer : la section « Select options »	171
Figure 19 : CR Viewer : la section « Display options »	172
Figure 20 : CR Viewer : l'affichage du résultat de « stats »	173
Figure 21 : CR Viewer : les trois textes représentés par « statsBar »	174
Figure 22 : CR Viewer : l'affichage du résultat de « frequencis »	174
Figure 23 : CR Viewer : diagramme circulaire représenté par « pie »	175
Figure 24 : CR Viewer : tous les maillons énumérés par « concordancer »	176
Figure 25 : Schéma des propriétés pour le texte français	216
Figure 26 : Schéma des propriétés pour le texte coréen	217
Figure 27 : Répartition des catégories grammaticales des maillons (<i>Fantôme</i>)	224
Figure 28 : Répartition des catégories grammaticales des maillons (<i>Arbre</i>)	225
Figure 29 : Sommaire du texte sur « Caniche », Wikipédia	253
Figure 30 : Exemple d'annotation : un bon nombre de singletons (<i>Fantôme</i>)	267

Règles de translittération

Nous suivons les règles de translittération proposées par l'Académie nationale de la langue coréenne.

Voyelles simples

ㅏ	ㅑ	ㅓ	ㅕ	ㅗ	ㅛ	ㅜ	ㅠ	ㅡ	ㅣ
a	eo	o	u	eu	i	ae	e	oe	wi

Diphthongues

ㅑ	ㅓ	ㅕ	ㅠ	ㅜ	ㅠ	ㅗ	ㅜ	ㅑ	ㅓ	ㅕ
ya	yeo	yo	yu	yae	ye	wa	wae	wo	we	ui

Consonnes occlusives

ㄱ	ㄲ	ㅋ	ㄷ	ㄸ	ㅌ	ㅍ	ㅃ	ㅍ
g/k	kk	k	d/t	tt	t	b/p	pp	p

Consonnes affriquées

ㅈ	ㅉ	ㅊ
j	jj	ch

Consonnes fricatives

ㅅ	ㅆ	ㅎ
s	ss	h

Consonnes nasales

ㄴ	ㅁ	ㅇ
n	m	ng

Consonne liquide

ㄹ
l

Introduction

Introduire les entités (personnes, objets et événements) et les reprendre en entretenant un enchaînement cohésif entre leurs expressions référentielles est un des éléments essentiels d'un texte bien formé. La notion de *chaîne de référence*, s'inscrivant à la croisée des problèmes de référence et de cohésion textuelle, représente la manière dont les expressions coréférentielles, appelées *mentions* ou *maillons*, se manifestent dans le texte et forment une chaîne d'éléments interdépendants.

Comment dès lors, pouvons-nous reconnaître la cohésion référentielle du texte alors que plusieurs formes différentes désignent une même entité comme illustré dans l'exemple ci-dessous ?

Record d'impopularité pour **Olaf Scholz** en Allemagne. **Le chancelier**, arrivé au pouvoir en décembre 2021, n'est plus soutenu que par un quart de l'opinion. Les défis s'accumulent pour **le successeur d'Angela Merkel**.

[...] La popularité d'**Olaf Scholz** et celle de **son** gouvernement sont en chute libre depuis mars et le début de la guerre en Ukraine. — (RFI, 22/08/2022).

Ici, les expressions désignant *Olaf Scholz* se présentent sous diverses formes comme le nom propre, le SN défini et le déterminant possessif. Comment un lien unissant ces différentes expressions renvoyant à un même référent, à savoir *les chaînes de référence*, se construit-il ? Quelles sont les caractéristiques de ce lien ? Le phénomène de la chaîne de référence se diversifie-t-il selon la langue cible ou selon le type de texte ? Toutes ces questions ont constitué le point de départ de notre thèse.

La notion de *chaîne de référence* est relativement récente. Bien que la notion elle-même soit introduite en 1975 sous le nom de *chaîne anaphorique* par le linguiste américain, Chastain, elle n'a guère été prise en compte dans le domaine linguistique.

Depuis les différents travaux de Corblin (1985, 1995), Charolles (1988), Givon (1983) et Schnedecker (1997), la relation entre les maillons a commencé à être traitée dans la

linguistique française. Dès lors, les chaînes de référence sont devenues un domaine de recherche important, non seulement pour la linguistique textuelle qui traite de la structure des textes et de leur interprétabilité mais aussi pour le domaine de la linguistique appliquée comme le Traitement Automatique des Langues (TAL). Depuis peu de temps, elles font l'objet de recherches comparatives, dans une perspective à la fois diachronique (ancien français vs. français moderne) et inter-langues. Ces approches sont développées au sein du projet ANR DEMOCRAT, entre autres, le projet dont nous nous inspirons et que nous présenterons au cours de cette thèse. Cependant, bien que le projet DEMOCRAT ait permis l'essor de nombreux travaux portant sur la manière dont la composition des chaînes de référence varie selon l'époque, le genre de discours ou la langue, le point de vue inter-langues se concentre principalement sur les langues romanes et germaniques, alors que langues asiatiques, notamment le coréen, sont négligées.

En outre, en linguistique coréenne, le thème de la référence a également fait l'objet de plusieurs travaux : les expressions référentielles ont été étudiées par plusieurs auteurs (Hong C. S, 2001 ; Chang K. H., 2002 ; Oh C. Y., 2012, entre autres), et des recherches comparatives sur les expressions référentielles, notamment les démonstratifs dans des langues distinctes (coréen-chinois, coréen-japonais et coréen-anglais) ont également été menées. Toutefois, la notion de *chaînes de référence* est un nouveau domaine pour les linguistes coréens : la description intégrée des expressions référentielles ou du phénomène lui-même sont encore absentes.

L'objectif de notre thèse est ainsi double. Le premier est de présenter la notion de *chaîne de référence* et les expressions référentielles des deux langues, français et coréen, et, le deuxième vise à analyser le phénomène dans un corpus bilingue de textes authentiques, d'un point de vue comparatif inter-langues.

Étant donné que l'époque (Perret, 2000 ; Capin, 2014 ; Glikman et *al.*, 2014), le genre textuel (Schneidecker, 2014) et la langue de rédaction (Baumer, 2015) sont les trois facteurs importants qui influencent les chaînes de référence, pour réaliser une comparaison efficace entre deux langues, nous avons décidé de sélectionner comme corpus un texte de chaque langue d'un même genre textuel de la même époque, le roman du début du 20^{ème} siècle. Ainsi choisis, le roman français *Le fantôme de l'Opéra* (Gaston Leroux, 1910) et le roman coréen intitulé *Arbre à feuillage persistant* (Shim Hoon, 1935) sont travaillés et analysés à l'aide de deux outils : SACR pour les annoter et CR Viewer pour réaliser et analyser le résultat de

l'annotation. De cette manière, nous espérons pouvoir montrer la manière dont les chaînes de référence se construisent et fonctionnent dans les romans français et coréen.

Notre thèse se constitue de trois grandes parties : d'abord, la description générale des chaînes de référence et des expressions référentielles, ensuite la méthodologie et les conventions d'annotation et, enfin la synthèse des analyses.

La première partie, qui a un caractère théorique, comprends quatre chapitres :

Dans le premier chapitre, quelques notions fondamentales autour de notre sujet sont discutées afin de poser le cadre théorique dans lequel la notion de *chaîne de référence* se situe. Nous commençons par présenter les notions de *cohérence* et de *cohésion* afin de mettre en évidence les conditions d'un texte bien formé. Puis, dans un second temps, les diverses approches sur *la référence* sont introduites pour approfondir notre compréhension du sujet. Dans un troisième temps, les notions d'*anaphore* et de *coréférence*, les notions connexes de *chaîne de référence* sont détaillées et comparées avec cette dernière.

Il sera question dans le deuxième chapitre des caractéristiques des chaînes de référence et de leurs constituants, les maillons. Plus particulièrement, nous présenterons le problème des bornes et les diverses modalités de cohabitation des chaînes de référence dans le texte. Nous expliquerons également les expressions référentielles susceptibles de s'employer comme maillon et l'influence du choix des maillons sur la nature des chaînes, ce qui nous amènera à travailler sur les expressions référentielles dans le chapitre 4. Ensuite, les trois facteurs qui concernent la variété de matériau lexical des chaînes, l'époque, le genre textuel et la langue de rédaction, seront explicités. Pour clore ce chapitre, les divers paramètres caractérisant les chaînes, qui serviront à l'établissement des stratégies d'analyse, seront détaillés.

Le troisième chapitre aura pour objectif, dans un premier temps, de présenter des recherches inter-langues portant sur les chaînes de référence en linguistique française et celles qui sont menées sur les expressions référentielles en linguistique coréenne. Dans un second temps, ce chapitre décrira quelques caractéristiques utiles de la langue coréenne telles que l'absence d'article et la présence d'un système alternatif indirect, qui pourront servir à la compréhension globale des disparités inter-langues.

Dans le quatrième chapitre, les divers emplois des expressions référentielles, les noms propres, les pronoms personnels et les syntagmes nominaux, seront présentés. Ces

descriptions aideront à fournir la base de la compréhension des propriétés des éléments constituant d'une chaîne en nous permettant de comparer les deux systèmes langagiers concernés.

La deuxième partie de notre thèse constituée de deux chapitres expose la méthodologie et nos choix d'annotation stratégique :

Le cinquième chapitre comprend la présentation du corpus, *Le fantôme de l'Opéra* et *Arbre à feuillage persistant*, et des outils choisis pour le travail d'annotation (SACR) et celui d'analyse des données (CR Viewer). Les raisons pour lesquelles nous avons sélectionné les deux textes romanesques seront détaillées et les points forts des outils seront décrits avec leurs modes d'emploi.

Dans le sixième chapitre qui clôt cette partie, nous décrirons nos conventions d'annotation. Inspirée du manuel d'annotation du projet DEMOCRAT et tenant compte des différences entre deux langues cibles, nous proposerons nos propres choix d'annotations.

La dernière partie qui est également le dernier chapitre de cette thèse sera consacrée à l'analyse des textes annotés. Les données fournies par l'outil CR Viewer seront interprétées à l'appui des exemples précis trouvés dans nos textes. Nous nous concentrerons particulièrement sur la composition des chaînes de référence dans chaque texte, à savoir la distribution des catégories grammaticales qui se manifeste différemment selon la langue. Sous cet angle, les conditions d'apparition de chacune des catégories grammaticales seront décrites. Par ailleurs, nous examinerons brièvement si la tendance observée dans notre corpus se présente également dans un autre type de texte, le texte encyclopédique, afin de voir, de façon générale, l'influence du genre textuel sur les chaînes de référence et la corrélation entre les traits de nos textes et les chiffres obtenus.

Cette thèse, qui est une « première » étude du domaine faite du point de vue comparatif entre le français et le coréen, permettra d'apporter de nouvelles données qui pourront enrichir les recherches sur les expressions référentielles et les chaînes de référence. Cela pourrait également être une opportunité d'introduire une nouvelle notion linguistique et d'apporter de nouveaux outils méthodologiques à la linguistique coréenne.

Première partie :

Description générale des chaînes de référence

Chapitre 1.

Problème de la cohérence et quelques notions fondamentales

Introduction

Ce chapitre a pour objectif de présenter les notions fondamentales qui seront exploitées dans notre thèse. Avant d'entrer dans le vif du sujet de cette recherche, nous ferons un bilan des notions de base telles que *la cohérence, la cohésion, la référence, l'anaphore, la coréférence et les chaînes de référence*.

Nous évoquerons, dans un premier temps (§1), le cadre linguistique dans lequel les chaînes sont exploitées. Nous commencerons par présenter le concept de *cohérence* et celui de *cohésion* afin de comprendre les conditions d'un texte cohérent (§1.1). Ensuite, nous introduirons, en renvoyant principalement à Charolles (1988), quatre plans d'organisation textuelle et montrerons la façon dont les différents éléments participent à la structuration d'un texte bien formé (§1.2).

Dans un deuxième temps (§2), nous présenterons les travaux linguistiques sur la référence. Nous expliciterons la manière dont la référence est devenue un sujet important dans la linguistique (§2.1) avant de définir la notion elle-même et ses notions connexes comme *le sens, la dénotation et la saillance* (§2.2). Puis, nous ferons un bref tour d'horizon des différentes approches théoriques sur le sujet (§2.3).

Dans un troisième temps, nous nous proposerons de distinguer les trois notions essentielles pour saisir l'objet de notre recherche : *anaphore, coréférence et chaîne de référence* (§3). Les différences entre les deux premiers termes seront précisées selon deux critères : l'autonomie référentielle et la relation symétrique des expressions référentielles (§3.1). Ensuite, la notion de *chaînes de référence* sera introduite et pour clore ce sous-thème, elle sera comparée avec d'autres relations des expressions référentielles qui sont l'anaphore et la coréférence (§3.2).

Ces travaux offriront une vue globale sur le sujet, à savoir les chaînes de référence, et établiront une base théorique pour ce qui va suivre tout au long de cette recherche.

1 Problème de la cohérence et de la cohésion

Un texte n'est pas qu'une suite de phrases alignées. Il ne s'agit pas non plus d'un simple ensemble d'idées thématiquement liées. L'objectif de cette première partie est de fournir une perspective globale sur la structuration textuelle. Pour ce faire, nous prendrons d'abord en compte les notions de *cohérence* et de *cohésion* proposées par de nombreux linguistes. Ensuite, la structure sous-jacente du texte, fondées sur les différents plans d'organisation textuelle, sera analysée avec un exemple précis.

1.1 La cohérence et la cohésion

Dans un texte, il y a « une quelconque unité structurale au-dessus de la phrase¹ » (Hallyday et Hasan, 1976). Regardons l'exemple suivant, un écrit d'enfant, que Schnedecker (1997 : 5) reprend des travaux de Reichler-Béguelin *et al.* (1988) :

(1) Sortie à la dune

Nous sommes partis de l'école à bicyclette pour aller à la côte découvrir la dune. L'euphorbe est une plante toxique, elle fait gonfler la langue et étouffe. Les plantes les plus répandues sont l'oyat et le chiendent. Les plus vieilles dunes ont plus de cinq mille ans. La criste marine est une plante comestible. Madame l'inspectrice accompagnait M. Berger. Nous avons continué la visite avec eux. Puis nous sommes rentrés à l'école. (Exemple de Turco, cité par Reichler-Béguelin *et al.*, 1988 : 132, in Schnedecker : 1997).

Les phrases thématiquement cohérentes ont beau être juxtaposées dans un même espace textuel, celui-ci semble se réduire à une suite de phrases sans aucun rapport les unes avec les autres, notamment en l'absence de liens référentiels attachant les phrases l'une à l'autre.

Pour qu'une suite de phrases constitue un texte bien formé, un certain nombre de critères ont été mis au jour dans le domaine de la linguistique textuelle. La cohésion et la cohérence comptent parmi ces critères. Salles (2006) présente le consensus des linguistes (Moeschler

¹ Traduit et cité par Charolles (2001 : 2).

& Reboul, 1994 ; Schnedecker, 1997, entre autres) concernant les deux notions en reprenant la proposition de Charolles (1988)². La cohérence est une notion liée à « l'interprétabilité du texte » tandis que la cohésion concerne « les moyens linguistiques » (Salles, 2006 : 1) comme la répétition, l'ellipse, l'anaphore, le connecteur, *etc.* qui contribuent à la continuité du texte :

[...] tout le monde est à peu près d'accord pour opposer d'un côté la cohérence, qui a à voir avec l'interprétabilité des textes, et, de l'autre les marques de relation entre énoncés ou constituants d'énoncés. Concernant ces marques, depuis M.A.K. Halliday et R. Hasan (1976), on tend à les regrouper sous le nom générique de cohésion. (Charolles 1988 : 53).

Prenant part à la continuité textuelle³ et à l'interprétabilité du texte, ces deux notions sont les concepts essentiels en linguistique textuelle. Reboul (1997 : 298) souligne que « la cohérence est souvent considérée comme l'équivalent pour le texte de ce qu'est la grammaticalité pour la phrase ». Comment établit-on alors des liens de cohérence et de cohésion dans un texte ?

De nombreuses recherches ont porté sur cette question et la première, chronologiquement parlant, est proposée par Charolles (1978 : 12), qui présente « les 4 méta-règles de cohérence » comme suit :

- méta-règle de répétition : « pour qu'un texte soit (microstructurellement ou macrostructurellement) cohérent, il faut qu'il comporte dans son développement linéaire des éléments à récurrence stricte. » (*ibid.* : 14) ;
- méta-règle de progression : « pour qu'un texte soit microstructurellement ou macrostructurellement cohérent, il faut que son développement s'accompagne d'un apport sémantique constamment renouvelé. » (*ibid.* : 20) ;
- méta-règle de non-contradiction : « pour qu'un texte soit microstructurellement ou macrostructurellement cohérent, il faut que son développement n'introduise aucun

² Charolles, M. (1988), « Les études sur la cohérence et la connexité textuelles depuis la fin des années 1960 », *Modèles linguistiques*, tome X, fasc.2, n° 20, 45-66.

³ Adam (2005 : 33) présente le terme de *liage* qui peut être employé comme l'équivalent de la continuité textuelle. Les unités textuelles sont « reliées entre elles ». Ces unités sont combinables avec d'autres constituants du texte soient qu'ils se trouvent proches, soient qu'ils sont à longue distance. Et les opérations sont effectuées par plusieurs étapes et plusieurs façons selon la situation dans la mesure où elles respectent la continuité textuelle.

élément sémantique contredisant un contenu posé ou présupposé par une occurrence antérieure ou déductible de celle-ci par inférence. » (*ibid.* : 22) ;

- méta-règle de relation : « pour qu'une séquence ou qu'un texte soient cohérents, il faut que les faits qu'ils dénotent dans le monde représenté soient reliés. [...] pour qu'une séquence soit admise comme cohérente il est nécessaire que les actions, états ou événements qu'elle dénote soient perçus comme congruents dans le type de monde reconnu par celui qui l'évalue. »⁴ (*ibid.* : 31).

Dans un texte cohérent, autrement dit, dans un texte qui respecte bien les conditions mentionnées ci-dessus⁵, certains dispositifs linguistiques comme les connecteurs, les différentes formes d'anaphores (et, plus largement, les expressions référentielles), les expressions introductrices de cadres de discours et les marques configurationnelles contribuent à la cohésion (Charolles, 1995a)⁶.

⁴ Charolles mentionne la notion de *relation de congruence* qui désigne des relations comme la cause, la condition, la conséquence pertinente entre les faits. Il admet qu'il s'agit d'« une relation assez lâche » comme elle se base uniquement sur « l'aperception d'un rapport de faits » (1978 :31). Dans les travaux ultérieurs, l'auteur présente des relations cause à effet, relation de contiguïté et relation de ressemblance. (Charolles : 1995a, 2011).

⁵ Salles (2006 : 2) signale que les règles de Charolles sont nommées différemment chez Reinhart (1980). Reinhart (1980) nomme trois des quatre règles de Charolles sous les noms de « cohésion, consistance et pertinence⁵ ». Ces trois notions sont comparables respectivement à la *répétition*, à la *non-contradiction* et à la *relation* telles que définies par Charolles :

[...] pour qu'un texte soit (globalement) cohérent, il doit répondre à chacun des trois critères suivants : cohésion, consistance et pertinence. Le premier est la condition de la concaténation linéaire des phrases dans le texte. Il demande que les phrases du texte soient formellement connectées. [...] Le deuxième est la condition exigeant que chaque phrase soit consistante avec la phrase précédente, à savoir, elles peuvent toutes être vraies dans le même état des choses (considérant notre supposition commune vers le monde).

[...] Les conditions de la pertinence est la moins appréhendée parmi les conditions pour la cohérence. Elles consistent en conditions sémantique et pragmatique et à la différence des deux autres précédentes, elles restreignent non seulement les relations entre les phrases du texte mais aussi les relations de celles-ci avec le thème sous-jacent du discours ainsi que leurs relations avec le contexte d'occurrence. (Reinhart, 1980 : 164, traduit par nos soins)

⁶ Charolles (1995a : 4) présente les quatre marques représentatives contribuant à la cohésion :

- Les connecteurs relient les phrases l'une de l'autre en indiquant des relations fonctionnelles entre elles.
- Les différentes formes d'anaphores « assurent des solidarités référentielles » entre les entités textuelles et

Dans ses travaux ultérieurs, Charolles (1995a ; 1997), en admettant toujours l'importance de ces marques de cohésion, ajoute « l'inférence de liaison » comme facteur fondamental de la cohérence. Cette idée, qui a été traitée par de nombreux auteurs⁷ à partir des années 1970, est un point qui mérite d'être reconnu au vu de l'existence des exemples⁸ où l'on ne trouve aucun des dispositifs linguistiques mentionnés comme marques de cohésion, mais qui restent toujours cohérents. Cependant, ce sujet étant susceptible de nous amener aux confins de la pragmatique, nous allons pour l'instant nous restreindre à la théorie de Charolles mentionnée dans les paragraphes précédents.

1.2 Quatre plans d'organisation textuelle

Un texte bien formé répond à certains critères d'organisation. Charolles (1988) distingue quatre plans d'organisation textuelle : la période, la chaîne, la portée et la séquence. Les marques d'organisation textuelle participent à « structurer la matière du discours, mais à des plans différents » (*ibid.* : 6) :

- « La période est une unité d'énonciation dont les membres ou composants (phrastiques) entretiennent des rapports de dépendance » (*ibid.*). Les phrases sont liées l'une à l'autre par des différents types de conjonctions : conjonctions de coordination⁹, de subordination¹⁰,

suscitent la naissance des chaînes de référence.

- Les expressions introductrices de cadres de discours ont la fonction de délimiter des domaines ou cadres du texte.
- Les marques configurationnelles ainsi que des alinéas et des organisateurs méta-discursifs, délimitent le continuum textuel.

⁷ Clark (1977) a traité cette notion sous le nom de « inférences de pontage ».

⁸ L'exemple que Charolles (1995a ; 2006) reprend à G. Brown & G. Yule (1983) qui montre qu'il n'y a aucune des dispositifs linguistiques entre A et B, mais qu'ils sont cohérents :

A : On sonne.

B : Je suis dans mon bain.

⁹ « Mais », « ou », « et », « donc », « or », « ni », « car » sont traditionnellement classés comme des conjonctions de coordination. Elles relient « deux éléments (termes, groupes, propositions) placés sur le même plan syntaxique » (Denis et Sancier-Chateau, 1994 : 129).

¹⁰ La conjonction de subordination (*que, puisque, quand, même si, pour que, etc.*) relie « deux éléments situés sur des plans syntaxiques différents » et « intègre dans la phrase la proposition qu'elle introduit et matérialise ainsi le

de juxtaposition¹¹, etc. Voici les exemples avec différentes conjonctions :

(2) a. J'écris ma thèse toute la journée **et** je sors le soir. (coordination)

b. Je crois **qu'**il viendra ce soir. (subordination)

c. Il n'est pas venu hier soir : il a une thèse à écrire. (juxtaposition)

- « Les chaînes sont constituées par des suites d'expressions coréférentielles » (*ibid.* : 8). Les diverses expressions référentielles comme les noms propres, les pronoms, les syntagmes nominaux, etc. qui renvoient au même référent peuvent composer une chaîne tout au long du texte : dans l'exemple suivant, les expressions référentielles renvoyant à « Elizabeth II » constituent une chaîne (mises en gras dans le texte).

(3) Quand **Elizabeth II** a fait **son** apparition sous la marquise dressée sur la pelouse du Haut-Commissariat britannique, l'invité s'est figé. **La souveraine** est plus petite qu'on ne l'imagine. — (*Le Monde*, 08/09/2022)¹²

- Les portées sont définies comme « toute portion de texte dont l'interprétation est indexée comme devant s'effectuer dans un certain cadre (ou espace) de vérité » (*ibid.* : 9). Dans un texte, certaines expressions linguistiques fonctionnent comme un cadre. Les expressions comme « Selon A », « D'après A » ou « Pour A » encadrent le contexte en attribuant « des propos à autrui », tandis que d'autres marqueurs d'univers de discours tels que « En France », « En 2019 » fournissent le cadre spatial ou temporel de ce qui vient d'être dit ou ce qui va être dit.

(4) **Selon nos informations**, le chef de l'Etat a décidé d'une vaste restructuration de son équipe de communication. — (*Les Echos*, 05/09/2018)¹³

lien de dépendance, la hiérarchisation entre propositions » (Denis et Sancier-Chateau, 1994 : 129-130).

¹¹ La grammaire traditionnelle définit la juxtaposition comme suit : « On appelle juxtaposition le mode de construction qui consiste à placer l'un à côté de l'autre, sans mot de liaison matérialisant le type de la relation, deux ou plusieurs termes mis sur le même plan syntaxique, c'est-à-dire occupant la même fonction au sein de la phrase ou de la proposition. » (Denis et Sancier-Chateau, 1994 : 312).

¹² « Elizabeth II : après soixante-dix ans d'un règne hors du commun, la souveraine laisse une empreinte durable sur la monarchie »

https://www.lemonde.fr/disparitions/article/2022/09/08/la-reine-elizabeth-ii-est-morte-apres-soixante-dix-ans-d-un-regne-hors-du-commun_6140784_3382.html

¹³ « Macron restructure la communication de l'Elysée »

- Les séquences sont considérées comme « le résultat du découpage du matériau discursif ». En d'autres termes, elles concernent la façon dont l'auteur « fractionne son texte » (*ibid.* : 9). À l'écrit, elles correspondent souvent au découpage en paragraphes, le locuteur indiquant la démarcation de son propos par la structure du texte. Dans l'exemple suivant, nous observons deux séquences, en d'autres mots, deux paragraphes différents.

(5) Les femmes continuent à être sous-représentées parmi les créateurs d'entreprises en France et ne représentaient en 2018 que 35 % d'entre eux, souligne l'Insee dans une étude publiée jeudi.

Ce pourcentage n'a guère évolué au fil des années, puisque seulement 32 % des créateurs d'entreprises étaient des femmes en 2010 et 33 % en 2014. — (*Le Point*, 15/09/2022)¹⁴

Ces différents plans d'organisation du discours peuvent cohabiter dans le texte. Autrement dit, ils interagissent au fur et à mesure que le texte progresse. L'exemple qui suit illustre la manière dont les quatre plans d'organisation textuelle coexistent dans un texte et participent à la structuration de ce texte :

(6) [**L'écrivain français Milan Kundera**, 90 ans, vient de retrouver **la nationalité tchèque QU'il** avait perdue en 1979.] [La décision a été prise le 6 novembre, MAIS le certificat qui l'officialise **lui** a été remis de la main à la main par l'ambassadeur de la République tchèque en France, **jeudi 28 novembre, LORS d'une cérémonie privée au domicile parisien du romancier.**]

L'écrivain est né à Brno, la deuxième ville du pays, en avril 1929. [**Il** avait été exclu du Parti communiste en 1970 ET **avait** été autorisé légalement à rejoindre la France en 1975, pour enseigner à l'université de Rennes.] [**Son** visa avait été renouvelé au bout de deux ans. MAIS en 1979, après la publication d'extraits du *Livre du rire et de l'oubli* dans *Le Nouvel Observateur*, puis d'un entretien au journal *Le Monde*, **il** avait été déchu de **sa nationalité tchèque.**] En juillet 1981, juste après l'élection de

<https://www.lesechos.fr/politique-societe/emmanuel-macron-president/macron-restructure-la-communication-de-lelysee-138129>

¹⁴ « Les femmes toujours sous-représentées parmi les créateurs d'entreprise, selon l'Insee »

https://www.lepoint.fr/societe/les-femmes-toujours-sous-representees-parmi-les-createurs-d-entreprise-selon-l-insee-15-09-2022-2490139_23.php

largement de Charolles (2002) a pour vocation de faire un rapide état des lieux des recherches linguistiques portant sur la *référence*. Par ailleurs, les notions-clés du domaine seront présentées avant de mettre en relief la problématique générale de cette étude.

2.1 Importance et aspects du problème

La *référence* est souvent définie comme le rapport entre une expression linguistique et un objet qui se situe en dehors du langage, à savoir, dans le monde extra-linguistique.

Ce concept est couramment traité par de grands philosophes occidentaux comme Platon, S. Mill, G. Frege, B. Russell, P. Strawson, W. Quine, *etc.* Cependant, d'après Charolles (2002), les linguistes ont essayé de contourner ce sujet en plaçant la référence à l'extérieur au champ des études. Pour Saussure, entre autres, le référent n'a pas de place dans la définition du signe linguistique dont les constituants sont le *signifiant* (image acoustique) et le *signifié* (concept)¹⁷. Par ailleurs, pour les structuralistes, la référence est un « phénomène de parole¹⁸ » qu'il ne convient pas de traiter en linguistique dont seule la langue fait partie.

Jusque dans les années soixante-dix, cette conception conserve une place prépondérante dans le domaine de la linguistique : la notion de *référence* est présentée dans un cadre restreint où elle n'est considérée que comme la fonction des signes linguistiques renvoyant à la réalité extra-linguistique. Kleiber (1981 : 11) présente trois définitions de ses prédécesseurs :

A : La référence est la fonction par laquelle un signe linguistique renvoie à un objet du monde extra-linguistique, réel ou imaginaire. (J. Dubois *et al.*, 1972 : 414)

B : Par référence, on entend généralement la fonction référentielle du message qui permet à un signe linguistique de désigner des êtres et des objets de l'univers extra-linguistique. (B. Pottier *et al.*, 1973 : 439)

C : La communication linguistique ayant souvent pour objet la réalité extra-linguistique, les locuteurs doivent pouvoir désigner les objets qui la constituent ;

¹⁷ Nous reprenons ce que Charolles (2002 : 10) cite : « Le signe linguistique unit non une chose et un nom, mais un concept et une image acoustique » (Saussure dans son *Cours de linguistique générale*, 1915/1968 : 158-159).

¹⁸ Saussure distingue la langue de la parole. La langue désigne l'ensemble des normes et l'outil pour la communication. En revanche, la parole signifie l'utilisation concrète et momentanée de la langue.

c'est la fonction référentielle du langage (le ou les objets désignés par une expression forment son référent). (O. Ducrot et T. Todorov, 1972 : 317)

Toutes ces définitions s'accordent en ce sens qu'elles ne dépassent pas la définition binaire du signe linguistique. Pourtant, ce point de vue ne fournit pas une explication précise et suffisante sur la manière dont le langage parle de la réalité. Le point de départ de notre discussion est donc de savoir, d'une part, si la référence devrait être incluse ou bien exclue des analyses linguistiques, et d'autre part, quelle approche il serait préférable de choisir pour bien l'appréhender dans le cadre de notre sujet d'étude.

Avec l'intérêt croissant porté à la linguistique du discours et les modèles cognitifs ou pragmatiques censés rendre compte des divers phénomènes linguistiques, les linguistes ont commencé à traiter le problème de la référence sous divers angles depuis quelques décennies. Avant d'évoquer ces diverses approches, nous allons procéder à un bref état des lieux de la notion de *référence* et de ses notions connexes car, comme le souligne Kleiber (1981), c'est un sujet très vaste qui est susceptible de susciter l'ambiguïté malgré diverses définitions proposées par plusieurs auteurs :

L'emploi fluctuant du verbe *référer* et, par conséquent, de l'expression *référence*, peut s'expliquer par l'existence d'une multitude de termes voisins comme *dénoter* (*dénotation* et *dénotatum*), *désigner* (*désignation* et *désignatum*), *dénommer* et *dénomination*, *identifier*, *renvoyer à*, etc., qui tantôt sont synonymes de *référer*, tantôt s'en différencient, la plupart de ces termes (cf. surtout *dénoter* et *dénotation*, que Geach, 1962, propose d'abandonner, tellement ils lui paraissent équivoques) n'échappant pas eux-mêmes à l'ambiguïté. Les traducteurs français de Frege et de Russell emploient les termes de *dénotation* et de *dénoter*, mais on pourrait tout aussi bien utiliser ceux de *référence* et de *référer*, et parler ainsi de *périphrases référentielles* au lieu de *périphrases dénotantes*. Chez Donnellan (1971), qui oppose *dénoter* à *référer*, certains syntagmes peuvent dénoter sans référer, mais non référer sans dénoter, alors que les SN référentiels de Heringer (1969), par contre, peuvent référer sans dénoter. (Kleiber, 1981 : 12)

Tenant compte du fait que plusieurs termes sont employés pour désigner la référence, notre tentative d'étudier les notions-clés autour de *référence* fournira une base théorique à ce que nous allons traiter tout au long de cette étude.

2.2 La notion de *référence*

Dans cette section, nous examinerons la notion de *référence*, ainsi que celles d'*existence* et de *sens*. Nous présenterons également la corrélation qu'entretiennent ces concepts, particulièrement utile pour la compréhension du problème de la référence.

2.2.1 Référence et existence

La *référence* concerne « la relation qui unit une expression linguistique au 'quelque chose' qu'elle exprime » (Kleiber, 1997 : 9). Depuis longtemps, dans la littérature philosophique et linguistique, les auteurs n'ont pas réussi à adopter une vision convergente sur la limite et la caractéristique de ce « quelque chose » qui est le référent. Nous allons donc clarifier les notions de *référence* et de *référent* du point de vue linguistique, notions essentielles pour l'interprétation du sens dans le texte.

Tout d'abord, une première question porte sur l'existence du référent. Comme le souligne Searle (1972 : 121, cité par Kleiber, 1997), le problème de la référence concerne nécessairement l'idée d'existence. Pour s'en rendre compte, il suffit de considérer l'exemple suivant :

(7) Nous n'avons pour l'instant reçu **aucune réponse**.

Dans l'exemple (7), « aucune réponse » n'a pas d'existence. De ce fait, on ne peut pas déterminer un « corrélat extensionnel » (Charolles, 2002 : 33) qui lui corresponde. Voici un autre exemple qui fait allusion à un point important de la relation de référence et d'existence :

(8) Delphine adore **les licornes**.

Dans (8), « la licorne » est une entité qui n'a pas d'existence dans la réalité. Pourtant, le référent vient immédiatement à l'esprit et permet de construire une représentation mentale de la licorne. À partir de là, une question essentielle se pose sur la notion d'*existence* : le référent doit-il forcément être une entité du monde réel ? En d'autres termes, comment peut-on considérer des entités abstraites ou imaginaires ?

L'approche classique de la *référence* souligne l'engagement ontologique, c'est-à-dire que les référents doivent exister dans le monde réel, indépendamment du langage. Cependant, d'après Kleiber (1997), cette conception objectiviste prête à controverse lorsqu'on rencontre les référents fictifs ou imaginaires comme une licorne, un dragon, Harry Potter, *etc.*, qui sont

effectivement considérés comme ayant un référent même si celui-ci n'est pas « tangible » dans la réalité. Par conséquent, il serait nécessaire d'« étendre l'accès de la référence à des mondes ou univers autres que le seul monde réel et donc d'élargir la définition initiale pour y intégrer ces mondes possibles » (Kleiber, 1997 : 11).

Le dogme de l'objectivisme amène certains auteurs au constructivisme en mettant en avant le fait que le monde n'est qu'un univers construit. Cette proposition semble raisonnable à certains égards car on ne voit pas le monde comme il est. Autrement dit, il y a toujours l'intervention de la perception de chacun et de sa propre conceptualisation lorsqu'on appréhende le monde. Cependant, Kleiber (1997) introduit deux points faibles du paradigme constructiviste. Premièrement, pour emprunter son expression, « ce monde perçu, conceptualisé, est ce que nous tenons pour la réalité » (*ibid.* : 13). Il est donc inutile, voire impossible de distinguer le monde réel du monde construit. Ensuite, la conception constructiviste ne peut pas répondre à la question de savoir comment une conceptualisation ou une modélisation du monde peut rester objective. Si chaque individu conceptualise le monde avec sa propre perception, le modèle mental du monde varie selon les personnes. Pourtant, chaque modèle paraît identique et on arrive à communiquer avec une « stabilité intersubjective » (*ibid.*). À partir de ces deux points, nous pouvons dire que le monde, au moins pour une part, existe réellement.

La primauté du monde réel devient donc un sujet important dans ce débat. Lorsqu'on réfère aux entités fictives ou imaginaires, c'est à partir de ce monde réel qu'on trouve des indices de sur l'entité dont il s'agit : comme le dit Kleiber (1997 : 15), « l'accès aux mondes possibles passe nécessairement par le monde réel ». De même, Larsson (1997, cité par Kleiber, *ibid.*), en mettant l'accent sur la prédominance du monde réel, souligne « la possibilité d'exprimer plusieurs modes d'existence différents » dans la langue (Larsson, 1997 : 55). Ainsi, une expression peut renvoyer à des entités réelles, fictives, hypothétiques, *etc.* Similairement, Charolles (2002) souligne, par l'explication claire et simple qui suit, que la condition nécessaire d'une référence est que l'expression doit être associable à ce qu'on conçoit comme une entité possible :

La seule condition requise pour qu'une expression réfère est donc qu'elle permette de fixer une entité dans la représentation mentale des états de choses dénotés par l'énoncé. (Charolles, 2002 : 36)

En suivant l'idée proposée par Kleiber (1997), Larsson (1997) et Charolles (2002) qui arrangent des approches objectivistes et constructivistes, dans les pages suivantes, nous

aborderons les notions essentielles à la compréhension de la *référence*.

2.2.2 Référence et sens

Comme mentionné précédemment (voir la section 2.1), le problème de la référence n'a pas trouvé de place dans la linguistique française d'inspiration saussurienne, et est resté donc en dehors du champ des études. Comment pouvons-nous alors aborder le sujet de manière linguistique ? À quels enjeux devons-nous faire face encore pour pouvoir aborder le sujet de cette thèse ? Dans les pages suivantes, nous apporterons les éléments de réponse à ce type de questions.

2.2.2.1 Le triangle sémiotique et la référence externe

La référence est représentée comme un des sommets du triangle sémiotique. Voici les deux triangles sémiotiques présentés par Kleiber (1997), reprenant celui d'Ogden et Richards (1923), ainsi que celui revisité par Ullmann (1952) :

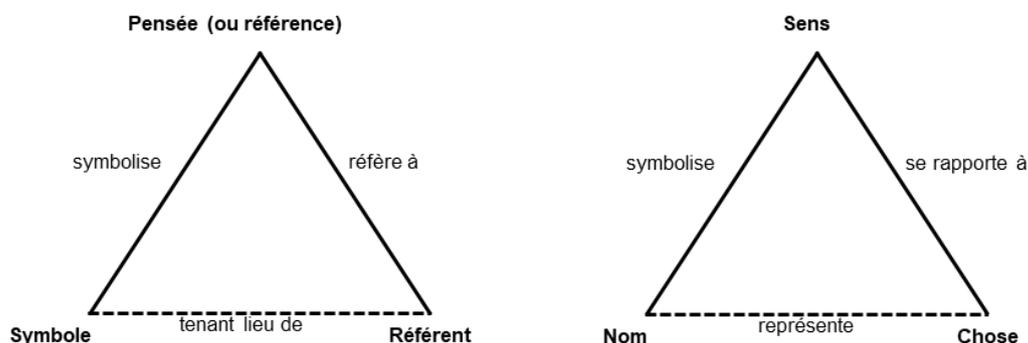


Figure 2 Triangles de C. Ogden & I. Richards (1923) et de S. Ullmann (1952), d'après Kleiber (1997)

Parmi les trois sommets du triangle, la branche linguistique est la branche gauche (*ibid* : 20). Dans cette optique, cette branche unit symboliquement ce qui est traditionnellement appelé dans la linguistique saussurienne « le signifiant » et « le signifié »¹⁹. Cette visualisation ternaire met en rapport ces deux notions avec un troisième élément externe, le réfèrent. Il en va de même pour la linguistique d'aujourd'hui comme le mentionne Kleiber (1997 : 16) : la

¹⁹ C'est ce qui est traditionnellement décrit chez les linguistes des années soixante-dix comme nous l'avons mentionné au début de cette partie (dans la section 2.1).

référence ne reste pas au niveau intra-linguistique, contrairement à ce qui est dit dans la perspective constructiviste. D'après lui, « le langage en tant que système de signes est tourné vers le dehors, vers ce qu'on appelle ou ce qu'on croit être la réalité, ou encore le monde, précisément parce qu'un signe n'est signe que s'il représente quelque chose d'autre que lui-même » (Kleiber, 1997 : 16). Par exemple, lorsqu'on dit « Albert Einstein », cette expression vise à désigner une entité hors du langage.

C'est à partir de cette idée qu'on commence à s'intéresser à la question de la référence, ainsi qu'à la nécessité de distinguer la référence du sens. Franckel (1998 : 61) souligne que la référence « relève d'un rapport de transparence et d'adéquation des unités de la langue aux idées qu'elles permettent de représenter, et qui sont elles-mêmes propres à constituer des représentations du monde ». De ce fait, « la référence de ces unités se confond avec leur sens ».

2.2.2.2 Sens, dénotation et référence

La notion de *référence* est souvent traitée en rapport avec celle de *dénotation*. Ces deux notions méritent d'être comparées pour qu'on puisse mieux appréhender le principe de la référence, et pour que celui-ci soit appliqué dans les étapes ultérieures de cette étude. Grâce à l'œuvre de Frege (1892), *Sinn und Bedeutung* (« Sens et référence »²⁰) qui introduit la notion de *sens*, qu'il distingue de celle de *référent*, les concepts sont déjà familiers. Dans son travail, Frege définit le sens comme la manière dont le référent est donné, à savoir son mode de présentation :

[...] il est naturel d'associer à un signe (nom, groupe de mots, caractères), outre ce qu'il désigne et qu'on pourrait appeler sa dénotation, ce que je voudrais appeler le sens du signe, où est contenu le mode de donation de l'objet. (Frege, 1892 : 103)

Ces deux notions distinctes ont des dénominations diverses : « Sinn (sens) - Bedeutung (dénotation ou référence) » comme nommées dans l'œuvre de Frege, « dénotation - connotation » selon Stuart Mill (cité par Kleiber, 1997 : 20), et, plus généralement, elles sont appliquées au domaine linguistique sous les noms de « dénotation - référence » ou « référence virtuelle - référence actuelle ».

²⁰ Traduit de l'allemand et introduit par Claude Imbert (1971).

Malgré ces diverses dénominations, la plupart des auteurs s'accordent relativement sur la nécessité de séparer le sens et la référence. Cela est évident lorsqu'on rencontre différentes descriptions définies renvoyant à un même référent telles que :

(9) a. Le père de la théorie de la relativité est Albert Einstein.

b. Le lauréat du prix Nobel de physique en 1921 est Albert Einstein.

Même si les SN définis de (9a) et (9b) n'ont pas le même sens, ils renvoient à une même personne « Albert Einstein ». Cette interprétation ne serait pas possible si l'on ne distinguait pas le sens du SN et son référent.

Charolles (2002), en choisissant le terme « dénotation », le définit ainsi :

La dénotation d'une expression définit les conditions de satisfaction que doit remplir un être quelconque pour que l'on puisse le désigner à l'aide de cette expression. Elle ne désigne ni une occurrence particulière d'une chose, ni une instance singulière d'un événement, ni non plus une incarnation spécifique d'une propriété. (Charolles, 2002 : 8)

Il s'agit d'une notion qui renvoie à un ensemble de caractères concernant une expression, à savoir à son sens général dans le dictionnaire. Elle peut potentiellement désigner un être particulier, par exemple, le nom « chien » dénote toute entité possédant les attributs nécessaires pour être définie en tant que « chien ». La notion même n'a rien à voir avec un acte de référence d'un être particulier et c'est pourquoi elle est définie comme « l'ensemble de conditions caractérisant une unité lexicale » sous le nom de « référence virtuelle » dans le travail de Milner (1976 : 64).

D'après la description dichotomique de Milner (1976 : 64), qui distingue la référence virtuelle de la référence actuelle²¹, « une unité lexicale ne peut avoir de référence actuelle que si elle est employée ». Autrement dit, sans emploi, cette unité reste virtuelle, en tant que sens non réalisé.

²¹ Ces deux notions de Milner (1976) sont interprétées par Reboul (1991 : 125-126) comme suit : « Si l'on s'en tient aux séquences nominales, on constate qu'une séquence nominale a une référence, celle-ci étant déterminée d'après les propriétés linguistiques de la séquence nominale. Cet ensemble de propriétés linguistiques correspond à un ensemble de conditions qui décrit un *type* ou une *classe* de référence possible. Le référent d'une séquence nominale, l'objet qu'elle désigne, est sa *référence actuelle*, l'ensemble de conditions qui lui est attaché est sa *référence virtuelle*, autrement dit son sens lexical ».

Pour synthétiser ces points de vue, nous pouvons décrire la différence entre l'acte de référer et celui de dénoter de deux manières différentes. D'abord, au niveau sémantique, l'acte de dénoter concerne le sens général d'un nom, alors que l'acte de référer nous conduit à l'identification d'un exemplaire particulier de ce nom. En conséquence, nous observons une sorte de concrétisation au niveau du sens entre la dénotation et la référence (du sens plus large et général au sens plus étroit et spécifique).

Ensuite, si nous regardons la relation de ces deux notions, nous pouvons dire que la référence est un processus d'extension de la dénotation : la référence est toujours réalisée par le biais du sens dénotatif du nom et d'autres éléments dans la phrase comme le déterminant et le prédicat. Comme le dit Charolles (2002 : 10), la référence n'est pas « un trait attaché aux seules expressions ». Autrement dit, c'est en fonction de sa relation avec d'autres éléments de la phrase qu'une expression référentielle atteint son objet qui se trouve à l'extérieur du langage.

Dans cette étude, en raison de sa compréhension intuitive et immédiate, nous allons principalement utiliser le terme « sens » au lieu de « dénotation ». Dans les cas nécessaires, les notions de « la référence virtuelle » et « la référence actuelle » seront employées.

2.2.2.3 La saillance des référents

Après avoir présenté les notions essentielles relatives à la notion de *référence*, nous nous intéressons, dans cette présente partie, au problème de la saillance des référents. Selon *Le Grand Robert de la langue française*, le terme *saillant* est défini comme « ce qui est en évidence, en relief, ce qui ressort du contexte et s'impose à l'attention »²². À partir de cette définition, la *saillance* peut être saisie comme une notion qui concerne *quelque chose mis en relief* parmi d'autres dans *le contexte*. Elle suscite l'attention dans divers domaines de recherche tels que la psychologie, l'informatique, la science cognitive, et, de même, différentes sous-catégories de la linguistique : il s'agit d'un des sujets les plus souvent traités dans le domaine de la référence discursive et nous ne sommes pas la seule à nous y intéresser. Dans cette partie, nous nous concentrons sur la notion elle-même et la façon dont elle influence la production ou l'interprétation du texte.

Schnedecker (2011 : 21) souligne que, malgré l'importance de la notion et la fréquence

²² Nous reprenons la citation de Hou (2021 : 21).

de son emploi, il n'y en a pas de définition convenue dans le domaine de la linguistique. En effet, ce n'est qu'à partir des années quatre-vingts que divers auteurs cognitivistes commencent à parler de la saillance et que son aspect discursif entre sur la scène linguistique. D'après la synthèse de Hou (2021 : 22), la *saillance* est traitée, par les cognitivistes, sous diverses terminologies : « activation (Chafe, 1976), familiarité (Prince, 1981), topique (Reinhart, 1981), topicalité (Givón, 1983b), accessibilité (Ariel, 1990), caractère donné (Gundel *et al.*, 1993) et centre d'attention (Grosz *et al.*, 1995 ; Walker *et al.*, 1998) ».

Les termes de la psychologie de la Gestalt sont souvent mentionnés lors de la compréhension de la notion de *saillance* : *Figure* et *Fond*. La première notion fonctionne comme une sorte de variable alors que la deuxième joue un rôle de cadre référentiel. En d'autres termes, la *Figure* est le référent mis en avant dans le *Fond* qui équivaut au contexte. À partir de cette idée, nous abordons la notion de *saillance* dans le domaine linguistique. La *saillance* renvoie à la mise en avant d'une entité linguistique par rapport aux autres, dans un contexte donné. La description suivante de Hou (2021) reprenant une partie de la définition de Neveu (2011 : 311) détaille la notion linguistique de *saillance* :

Dans le domaine linguistique, la saillance se rapporte à une propriété des entités dans la représentation discursive, et elle sert plus particulièrement à « décrire le statut de centralité de certains référents dans la conscience des partenaires de l'énonciation. Un référent est saillant s'il s'impose à l'attention. [...] ». (Hou, 2021 : 21)

La *Grammaire Méthodique du Français*²³ (2018 : 965) insiste sur le fait qu'il est difficile d'ignorer un référent ayant une forte saillance dans une situation donnée. Les auteurs supposent une situation où le locuteur se plaint d'un trompettiste qui joue dans son environnement, « Il nous casse les oreilles ! ». Dans cette situation, on peut facilement trouver à qui réfère le pronom « il ». L'utilisation sans antécédent des pronoms permet de comprendre le mécanisme cognitif de *saillance*. Regardons l'exemple (10) :

(10) a. **Cindy** est luxembourgeoise. **Elle** apporte souvent des saucisses du Luxembourg.

b. **Il** arrive dans cinq minutes. (dit le secrétaire à un étudiant qui attend le directeur)

L'interprétation du pronom personnel « elle » de l'exemple (10a) s'appuie sur son antécédent « Cindy ». Dans ce cas, c'est à partir de ce qui est fourni par le texte qu'on identifie le référent. En revanche, dans l'exemple (10b), « il » est interprété grâce au haut degré de

²³ Désormais la *GMF*.

saillance considéré comme partagé dans la mémoire du locuteur et celle de son interlocuteur. Ainsi, l'interprétation de certaines expressions référentielles s'appuie sur le critère de la saillance déjà construit dans la mémoire discursive des participants de l'énonciation, ou bien, pour emprunter le terme de Brown et Yule, (1983), la représentation mentale²⁴.

Dans la linguistique d'aujourd'hui, la saillance est souvent considérée comme étant influencée par des facteurs variés. Pour reprendre le terme de Hou (2021 : 39), « un éventail de facteurs divers » contribue à la saillance d'une entité et entraîne « l'identification du *bon* référent ». En outre, Col (2012 : 7, cité par Hou, 2021) souligne également que des mécanismes généraux de la cognition humaine tels que la catégorisation, l'invariance perceptive, la compositionnalité gestaltiste, la composition séquentielle hiérarchique²⁵ sont des facteurs qui se répercutent sur la saillance.

La liste suivante proposée par Landragin (2012) montre le caractère multifactoriel de la notion. L'auteur distingue d'abord la dimension physique de la saillance de sa dimension cognitive. Puis, celle-ci est classée en trois sous-catégories :

1. Les facteurs physiques liés à la forme de l'énoncé : saillance intrinsèque au mot ; saillance due à une mise en avant explicite lors de l'énonciation ; saillance due à

²⁴ Hou (2021) souligne que la notion de *représentation mentale* se construit en fonction du déroulement de l'énonciation. Les locuteurs globalisent toutes les informations du contexte, de la situation de l'énonciation, des connaissances qu'ils ont sur le monde, etc. afin de suivre l'activité langagière qui a un caractère linéaire.

²⁵ L'auteur nous renvoie à Victorri (2010), l'article intitulé « Spécificité du langage : des questions à approfondir » (in Rapport de prospective du projet PIRSTEC) pour la compréhension des notions.

- La catégorisation : il s'agit du mécanisme cognitif qui consiste à « regrouper des inputs semblables, proches sur la plupart des dimensions perçues, pour les traiter de la même manière » (*ibid* : 1).
- L'invariance perceptive : elle joue une fonction complémentaire de la *catégorisation*, c'est-à-dire qu'elle identifie les mêmes inputs « qui peuvent être très éloignés sur certaines dimensions perçues, mais qui sont les transformations les uns des autres par des mouvements moteurs (translations ou homothéties dans l'espace pour le système visuel, translations dans la dimension fréquentielle pour un input sonore, etc.) » (*ibid* : 1-2).
- La compositionnalité gestaltiste : des stimuli complexes sont traités par « le principe gestaltiste d'interaction entre le tout et les parties » (*ibid* : 2). Chaque partie du tout participe à la compréhension du sens global de ce dernier et réciproquement, « c'est l'appartenance à ce tout qui permet de donner un sens à chacun des éléments » (*ibid* : 2).
- La composition séquentielle hiérarchique : elle traite « des éléments de niveaux hiérarchiques différents disposés linéairement » (*ibid* : 2).

une construction syntaxique dédiée ; saillance syntaxique liée à l'ordre et à la fréquence d'apparition des mots ; saillance liée aux fonctions grammaticales ; saillance indirecte par transfert grammatical de saillance.

2. Les facteurs physiques liés au sens de l'énoncé : saillance liée à la sémantique des mots ; saillance liée au rôle thématique ; saillance liée au thème et au topique de l'énoncé ; saillance liée au propos de la conversation ; saillance liée à des inférences ; saillance indirecte par transfert sémantique de saillance.

3. Les facteurs cognitifs : saillance liée à l'intention ; saillance liée à l'attention ; saillance liée à la mémoire à court terme ; saillance liée à la mémoire à long terme ; saillance liée à la personnalité ; saillance liée à l'affect et aux émotions.

(Landragin, 2012 : 29)

Les facteurs énumérés dans le premier groupe sont formels. En d'autres termes, la saillance de ce type « se caractérise par une trace physique détectable immédiatement dans le signal audio ou dans la suite de mots » (*ibid* : 25). Par rapport au premier groupe, les facteurs de la deuxième liste ont des traces physiques moins évidentes. La saillance est saisie par le sens de l'énoncé ou de l'énonciation. En outre, le dernier groupe se constitue de facteurs n'ayant pas de traces physiques immédiates. La saillance dont on parle ici dépend des mécanismes cognitifs.

Ainsi, la *saillance*, influencée par plusieurs facteurs, intervient dans la situation de communication linguistique afin d'identifier le bon référent. En fonction du statut de centralité ou de la prééminence du référent dans la conscience du locuteur et de son interlocuteur, la stratégie de la communication varie. Autrement dit, le degré de saillance a des répercussions sur le choix de l'expression référentielle (Ariel, 1990). Nous aurons l'occasion d'en reparler dans la section 2.3.2 *infra* et pour le moment, nous nous contentons d'avoir introduit la notion de *saillance* et ses caractéristiques.

2.3 Diverses approches linguistiques du problème de la référence

La notion de *référence*, souvent négligée dans le domaine de la linguistique jusqu'au début du 20^{ème} siècle²⁶, est devenue un des sujets importants dans la linguistique

²⁶ Voir les points 2.1 et 2.2.2.1 *supra*.

d'aujourd'hui. Notre étude est inspirée par le projet ANR DEMOCRAT : « Description et Modélisation des Chaînes de Références : outils pour l'Annotation de corpus (en diachronie et en langues comparées) et pour le Traitement automatique », qui traite ce problème et ses domaines connexes. Plus précisément, ce projet vise à fournir une description intégrale des expressions référentielles et des chaînes de référence *via* la construction et l'annotation du corpus²⁷. Plusieurs recherches se développent au sein de ce projet afin de caractériser les chaînes de référence y compris leur comportement textuel et leur typologie.

Dans cette partie, nous introduirons les trois types d'approches théoriques sur la référence proposés par le projet DEMOCRAT : les approches en sémantique grammaticale référentielle, les approches discursivo-fonctionnelles et les approches configurationnelles ou relationnelles (résumé du projet DEMOCRAT, 2015 : 4-5).

2.3.1 Les approches en sémantique grammaticale référentielle

Dans le domaine de la sémantique grammaticale référentielle, plusieurs auteurs ont travaillé respectivement sur diverses catégories d'expressions référentielles comme cités dans le projet DEMOCRAT (2015 : 4) : « les pronoms personnels (Kleiber, 1994), les pronoms démonstratifs (Cadiot, 1988 ; Corblin, 1987, 1998 ; Charolles, 1995b, 1997b ; Kleiber, 1994 ; Maillard, 1994 ; Sales, 2008) ou corrélés (l'un/l'autre, le premier/le second, Schnedecker, 2006), le SN démonstratif (Apothélos & Chanut, 1997 ; Béguelin, 1998 ; Corblin 1987 ; De Mulder, 1997, 2001, Schnedecker, 2006), le SN défini (Corblin, 1987 ; Kleiber, 1981, 1988), le SN possessif (Kleiber, 2009 ; Heinz, 2003), les anaphores zéros (Li & Thompson, 1979 ; Fillmore, 1986), les noms propres (Kleiber, 1981 ; Gary Prieur, 1994 ; Jonasson, 1994 ; Leroy, 2004) ».

En reprenant cette liste bibliographique récapitulée par DEMOCRAT, nous ajoutons également celles qui ne sont pas mentionnées dans le projet, mais qui méritent d'être incluses dans la liste : pronom personnel (Reboul, 1991), anaphore zéro (Zribi-Hertz, 1986), nom propre (Molino, 1982 ; Wilmet, 1995 ; Gary-Prieur, 2016 ; Gary-Prieur & Noailly, 2019). La plupart de ces travaux ont pour objectif principal de décrire chacune des catégories d'expressions référentielles.

²⁷ <http://www.agence-nationale-recherche.fr/Projet-ANR-15-CE38-0008>

2.3.2 Les approches discursivo-fonctionnelles

Les approches discursivo-fonctionnelles, appelées également *approches cognitivistes*, s'intéressent notamment aux problèmes liés à la *topicalité*. L'une des plus connues est la théorie d'Ariel (1990) portant sur l'accessibilité cognitive du référent, et s'appuyant sur la notion de *saillance*. Comme nous l'avons présenté dans la section 2.2.2.3 *supra*, la *saillance* est une notion qui « désigne avant tout l'émergence d'une figure sur un fond » (Landragin, 2011 : 68). Nous pouvons aussi dire que « un référent est saillant s'il s'impose à l'attention » pour reprendre la définition de Neveu (2000 :100). En résumé, un référent est saillant quand il est mis en avant parmi d'autres de son contexte.

Dans la théorie d'Ariel, le degré de proéminence des représentations mentales des référents diffère selon l'expression référentielle employée (voir la section 2.2.2.3 *supra*). Plus le référent est saillant, plus l'expression référentielle marquera un degré élevé d'accessibilité cognitive du référent. Dans le cas contraire, le locuteur utilisera un marqueur de basse accessibilité référentielle comme le nom propre qui contient déjà lui-même l'information suffisante sur le référent. Par exemple, le nom propre « Milan Kundera » de l'exemple (2) *supra* est un marqueur référentiel plus précis et plus direct que le pronom personnel « il » qui ne contient qu'une information partielle sur le référent. En termes plus précis, le pronom personnel « il », en tant que tel, n'indique que le genre, le nombre et la fonction grammaticale du référent : le référent en question est le sujet, masculin et singulier (nous aurons l'occasion d'y revenir).

La théorie du centrage (Grosz *et al.*, 1995 ; Walker *et al.*, 1998 ; Cornish, 2000, cités dans le projet DEMOCRAT) appartient également aux approches cognitivistes. D'après la citation de Cornish (2000 : 2), Walker *et al.* (1998 :1) indiquent que cette théorie a pour objet de modéliser « le centre d'attention des interlocuteurs au sein d'un discours [en termes de] rapport entre l'état attentionnel, de complexité inférentielle et de forme des expressions référentielles employées » (traduit dans Cornish, 2000). Cornish (2000 : 2-4) ajoute qu'elle s'intéresse au « degré de cohérence d'un segment de discours » et qu'on peut prévoir les référents qui sont « classés en fonction de leur topicalité ». Parmi les référents, le centre préféré (Cp) est le plus important dans un texte. Deux autres types de centres sont également présentés dans cette théorie : le centre anticipateur (Ca) et le centre rétroactif (Cr). Le premier désigne tous les centres dans un texte dont les interlocuteurs pensent qu'ils sont susceptibles de devenir le Cr qui est « une entité topique de discours local » du segment qui suit. Pour reprendre ce qui a été défini par Glikman, Guillot-Barbance et Orby (2014 : 45), cette théorie

tente de prévoir l'emploi des expressions référentielles en fonction d'une hiérarchie des référents.

2.3.3 Les approches configurationnelles ou relationnelles

Les approches dites « configurationnelles ou relationnelles » (termes proposés dans le cadre du projet DEMOCRAT, 2015 : 4) décrivent les liens entre une expression référentielle et sa source, à savoir la relation d'anaphore (Schneidecker *et al.*, 1994 ; Apothéloz, 1995 ; Kleiber, 2001). Cette notion concerne « une relation de dépendance orientée » (Corblin, 1985a : 178) entre deux expressions : la source, qui fonctionne en tant qu'indice de l'interprétation, et l'anaphorique, qui est une expression référentielle exigeant un indice de son contexte²⁸.

Dans cette approche, on se focalise sur les différents types d'anaphore. Celle-ci peut être classée sous plusieurs types en vertu de critères variés. Tout d'abord, suivant les catégories lexicales de l'anaphorique, les anaphores pronominales et nominales se distinguent les unes des autres (Milner, 1982 :18).

Les divers matériaux lexicaux de l'anaphore peuvent aussi fonctionner comme critère de classement. L'anaphore fidèle désigne le cas où l'on emploie le même matériau lexical pour l'antécédent et l'anaphorique (Apothéloz, 1995 : 36, entre autres). Dans le cas de l'anaphore infidèle, le nom employé pour l'antécédent et celui de l'anaphorique sont différents. Voici des exemples d'anaphore fidèle (11) et infidèle (12) :

(11) Mon voisin chante **une chanson** tous les jours. Maintenant, je connais **cette chanson** par cœur.

(12) J'ai croisé **un gros chien** dans la rue. **Cet animal** était très méchant.

Nous pouvons également distinguer l'anaphore coréférentielle et l'anaphore associative. Webber (1988 :4)²⁹ explique la différence de mécanisme entre ces deux anaphores par les figures qui suivent :

²⁸ On en reparlera dans la section 3.1 *infra* du point de vue comparative avec la notion de *coréférence*.

²⁹ Evoke, Specify and Co-Specify (Webber, 1988 : 4).

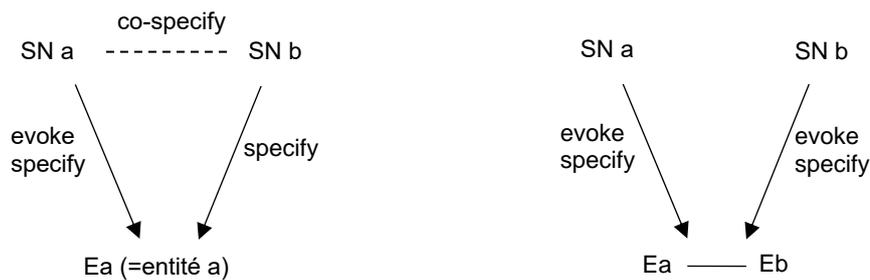


Figure 3 Anaphore coréférentielle et anaphore associative (Webber, 1988 : 4)

Dans la première figure, deux SN spécifient une même entité. Plus précisément, le « SNa » évoque une entité puis le « SNb » la spécifie encore une fois, et cela établit une relation de co-spécification entre les deux SN comme « un gros chien » et « cet animal » dans l'exemple (12) mentionné ci-dessus. En revanche, dans la deuxième figure, le « SNa » évoque et spécifie une entité (Ea) et le « SNb » le fait aussi pour une autre entité (Eb). Toutefois, ces deux entités évoquées par deux SN différents sont associées d'une certaine manière. Dans l'exemple (13) ci-après, « le marché » et « les légumes » ne réfèrent pas à la même entité. Cependant, étant donné que l'on trouve sans doute « les légumes » « au marché », nous pouvons appréhender facilement la relation associative entre « le marché » et « les légumes ».

(13) Je vais **au marché** tous les samedis matin. **Les légumes** sont plus frais qu'au supermarché.

Malgré le fait que ces trois axes de recherches aient été travaillés respectivement dans le domaine de la référence, ils servent tous ensemble de fondements théoriques pour notre objet d'étude, les chaînes de référence. Ils précisent les caractéristiques des constituants fondamentaux des chaînes de référence et les contraintes pesant sur leur emploi, les notions essentielles telles que la *saillance*, la *topicalité*, et enfin les différents types de relations entre les expressions référentielles. Toutes ces recherches ouvrent des horizons à une nouvelle facette du domaine de la référence.

3 Anaphore, coréférence et chaînes de référence

Nous venons de voir que les approches configurationnelles ou relationnelles dans le

domaine de la référence traitent le phénomène d'anaphore et s'intéressent aux divers sous-types de l'anaphore. Nous avons également remarqué que la *coréférence* et l'*anaphore* sont souvent placées sur le même plan lorsqu'on parle de la relation référentielle.

Cette partie vise à présenter ces deux notions de manière plus approfondie, et sera organisée de la façon suivante : dans un premier temps, la notion de *coréférence* et celle d'*anaphore* seront distinguées afin de servir de base, dans un second temps, nous reviendrons sur la compréhension du phénomène de *chaînes de référence*.

3.1 Anaphore et coréférence

Les notions d'*anaphore* et de *coréférence*, qui partent de la même catégorie, sont cependant distinctes, comme l'évoque la *GMF* (2018) :

Une expression linguistique qui entretient une relation de coréférence avec une expression antérieure n'est pas nécessairement anaphorique, dans la mesure où son interprétation référentielle n'en découle pas. (Riegel *et al.*, 2018 : 1030-1)

La distinction entre ces deux notions se fait à différents niveaux. La distinction la plus saillante entre l'anaphore et la coréférence porte sur le nombre de constituants auxquels elles réfèrent. La première ne s'applique qu'à une relation entre deux expressions :

Une expression est anaphorique si son interprétation référentielle dépend nécessairement d'une autre expression qui figure dans le contexte antérieur. (Riegel *et al.*, 2018 : 1029)

En comparaison, la coréférence peut désigner une relation liant plusieurs expressions. Les autres différences les plus saillantes entre l'anaphore et la coréférence sont l'autonomie référentielle d'expression et la relation symétrique des entités que nous évoquerons plus amplement dans les prochains points afin de faciliter la comparaison entre ces deux notions.

3.1.1 Autonomie référentielle

Certaines expressions référentielles sont autonomes, c'est-à-dire elles ne dépendent pas d'un autre élément de la phrase pour être interprétées, tandis que certaines autres exigent une information supplémentaire issue du contexte :

(14) Le prince Harry... Le fils cadet du prince Charles... Le petit frère du prince William...

Les trois expressions référentielles dans l'exemple (14) fonctionnent de façon autonome et se rapportent au même référent visé. Cet exemple reflète ce que Milner (1982) appelle la relation coréférentielle :

Il y a relation de coréférence entre deux unités référentielles A et B quand elles se trouvent avoir la même référence – ce qui peut arriver sans que l'interprétation de l'une soit affectée par l'interprétation de l'autre. (Milner, 1982 : 32)

Par contraste, la relation entre les expressions référentielles « non autonomes du point de vue de la référence » (Schnedecker, 2019 : 3) et qui dépendent fortement d'autres composants du texte pour être saturées constitue ce qu'on appelle une anaphore, une relation de dépendance interprétative entre deux expressions référentielles pour reprendre encore une fois la définition de Milner (1982) :

Il y a relation d'anaphore entre deux unités A et B quand l'interprétation de B dépend crucialement de l'existence de A, au point qu'on peut dire que l'unité B n'est interprétable que dans la mesure où elle reprend – entièrement ou partiellement – A. (*ibid.* : 18)

Dans l'exemple suivant, l'adjectif possessif « ses » et le pronom « il » ne peuvent pas être interprétés sans l'aide d'un autre composant du texte. Compte tenu du contexte, on comprend que ces expressions réfèrent respectivement à un personnage, « Alain », qui est appelé « antécédent »³⁰ dans la relation anaphorique :

(15) Quand au bout d'une semaine **Alain** revit **ses** copains dans un bistrot (ou chez Charles, je ne sais plus), **il** interrompit tout de suite leur bavardage... — (extrait du roman *La fête de l'insignifiance* par Milan Kundera, p.37).

3.1.2 Relation symétrique des entités

Nous nous intéressons désormais à la relation entre les entités dans les cas de coréférence et d'anaphore. La coréférence, qui s'appuie sur les deux expressions référentiellement autonomes, implique une relation symétrique entre ces deux mêmes

³⁰ « Ainsi définie, l'anaphore représente un phénomène de dépendance interprétative entre deux unités, la seconde ne pouvant recevoir un sens référentiel sans avoir été mise en connexion avec la première, laquelle est appelée, selon les terminologies, antécédent. » (Reichler-Béguelin, 1988 :17).

expressions, alors que l'anaphore comporte une relation asymétrique entre l'anaphorique et l'antécédent. Suivant la comparaison proposée par Corblin (1985a : 178), ces deux relations référentielles peuvent être schématisées ainsi :



Ce point peut être plus facilement compris en rapport avec la notion d'*autonomie référentielle*. Les deux expressions « le prince Harry » et « le fils cadet du prince Charles » de l'exemple (a) sont des expressions autonomes et donc interprétables indépendamment l'une de l'autre. La direction des flèches dans (a) montre aussi qu'il y a une relation symétrique entre ces expressions coréférentielles. Par contre, l'anaphore qui est, pour reprendre la définition de Corblin (1985a : 178), « une relation de dépendance orientée » vers l'antécédent, est représentée par (b). L'interprétation de l'anaphorique « il » dépend complètement de ce qui est précède, à savoir « le prince Harry ». L'orientation de l'interprétation est donc asymétrique.

Ainsi, ces deux notions, la *coréférence* et l'*anaphore*, se distinguent par deux aspects : le caractère de leurs constituants et le type de relation qu'ils établissent entre eux. À travers cette conception, nous aborderons une nouvelle notion : les *chaînes de référence*.

3.2 Chaînes de référence

Dans cette section, nous introduirons la notion de *chaîne de référence* qui est le sujet de notre étude. Cette notion, relativement récente dans la linguistique, sera examinée en comparaison avec celles d'*anaphore* et de *coréférence*.

3.2.1 La notion de *chaîne de référence*

Les *chaînes de référence* (abrégé CR) décrivent la cohésion référentielle entre les différentes entités linguistiques renvoyant à un même référent³¹. Il s'agit d'une notion

³¹ D'après Charolles (1987 : 33), dans le domaine de la constitution de CR, le référent peut être soit une entité du

essentielle de la linguistique, étroitement liée aux problèmes de la référence et de la cohésion textuelle. Cette notion a été introduite par C. Chastain (1975), dans son article intitulé « Reference and Context », sous le nom de « chaîne anaphorique » afin de désigner « une séquence d'expressions singulières apparaissant dans un contexte de sorte que si l'une d'entre elles se réfère à quelque chose, toutes les autres s'y réfèrent aussi³². » (*ibid* : 204-205, cité dans Corblin, 1995 : 16).

À partir des années 1980, la notion de *CR* a suscité beaucoup d'intérêt chez les linguistes. Elle a été redéfinie par F. Corblin (1985b : 123) comme « la suite des expressions d'un texte entre lesquelles l'interprétation construit une relation d'identité référentielle ». L'existence des *CR* prouve que les langues naturelles ont la capacité de « construire et ressaisir des objets de référence au fil d'un discours ou d'un texte » (*ibid* : 124).

L'exemple suivant est un extrait d'article du journal *Le Monde* évoquant le discours adressé par Bob Dylan au jury du prix Nobel :

(16) Fin du feuilleton **Bob Dylan**, prix Nobel de littérature. **Le chanteur américain** avait jusqu'au 10 juin pour transmettre à l'académie suédoise **son** discours de réception, faute de quoi la récompense de 8 millions de couronnes suédoises (837 000 euros) **lui** aurait échappé. La secrétaire perpétuelle de l'institution, Sara Danius, a annoncé, lundi 5 juin, que **le lauréat s'était exécuté**, en qualifiant le résultat, mis en ligne, d'« extraordinaire et, comme on pouvait s'y attendre, éloquent ». — (*Le Monde*, 08/06/2017)³³.

Comme le montre cet exemple, différentes expressions sont utilisées dans le texte pour désigner une même entité extralinguistique. Ici, « Bob Dylan » est désigné par les syntagmes nominaux (abrégé SN) « le chanteur américain » et « le lauréat » ainsi que par l'adjectif possessif « son » et le pronom « lui ». Dans le verbe pronominal « s'était exécuté », « se » réfère également au sujet de cette action qui est « le lauréat », c'est-à-dire « Bob Dylan »³⁴.

réel, soit une entité de la fiction, autrement dit, « l'existence matérielle » du référent n'est pas importante.

³² « ...a sequence of singular expressions occurring in a context, such that if one of them refers to something then all of the others also refer to it ».

³³ « Bob Dylan reçoit son prix Nobel sur une note d'Homère ».

http://www.lemonde.fr/musiques/article/2017/06/08/bob-dylan-recoit-son-prix-nobel-sur-une-note-d-homere_5140401_1654986.html

³⁴ Il se pourrait que l'inclusion du verbe pronominal dans la catégorie des expressions référentielles soit contestable.

Ces expressions référentielles sont appelées des « maillons ».

Schnedecker (1997 : 10) souligne, en empruntant le commentaire de Corblin (1987), que la notion de *CR* s'impose à partir du moment où il y a au moins trois maillons sur lesquelles on travaille :

Chaîne permet de dépasser les contextes de simple succession de deux termes auxquels se limite le plus souvent le linguiste qui sort du domaine phrastique. (Corblin, 1987 : 7)

En d'autres termes, comme l'indiquent Schnedecker et Landragin (2014 : 4), de simples « paires d'enchaînements référentiels », l'anaphore et la coréférence, ne sont pas suffisantes à rendre compte du phénomène référentiel sur le « long terme ». Nous choisissons de suivre cette idée afin d'examiner les spécificités des *CR* qui les rendent discernables avec d'autres phénomènes linguistiques.

3.2.2 Les différences entre anaphore, coréférence et chaîne de référence

Nous avons essayé de distinguer les notions de *coréférence* et d'*anaphore* dans la section 3.1 *supra*. Après avoir présenté la notion de *CR* à un rang supérieur à celui de ces deux conceptions, nous allons comparer ces trois notions.

Le tableau suivant est une partie de ce que Schnedecker (2019 : 14) récapitule sur les différences et la relation entre les trois notions que sont l'*anaphore*, la *coréférence* et les *CR* :

Par exemple, suivant Ariel (1990) qui inscrit les « \emptyset , réfléchis, traces QU- et accords » dans son échelle d'accessibilité référentielle, le pronom réfléchi est considéré comme une des expressions référentielles et est bien compris dans les chaînes.

En revanche, Schnedecker (2021 : 78) distingue les pronoms réfléchis ayant le fonctionnement référentiel de ceux qui ne l'ont pas. D'après cette distinction, seuls les pronoms réfléchis avec les fonctionnements « réfléchis » (ex. Chloé se lave les mains) et « réciproques » (ex. Lara et Guillaume se parlent dans le jardin) sont référentiels alors que le pronom « se » dans des constructions « anticausatives » (ex. La douleur s'estompe,) et « médiopassives » (ex. Ce plat se mange avec des baguettes) ne sont pas référentiels (Schnedecker reprennent à Barque *et al.* (2020) ces deux derniers exemples). Vu la complexité du phénomène, pour l'instant, nous nous contentons d'évoquer la contestabilité du sujet et nous en reparlerons dans les chapitres qui suivent.

	Anaphore	Coréférence	Chaîne de référence
Nombre d'expressions référentielles (ER)	2	2 ou n	$n \geq 3$
Nature des ER	Au moins une expression non autonome référentiellement	Expressions référentielles autonomes et ou non autonomes	
Relation référentielle	Identité et non identité	Identité	
Propriétés de la relation	Asymétrie, atransitivité	Symétrie, transitivité	Toutes
Domaines linguistiques couverts	Linguistique phrastique et interphrastique		Linguistique discursive & linguistique des genres

Tableau 1 Extrait d'un récapitulatif des différences entre anaphore, coréférence et CR (Schnecker, 2019 : 14-15)

Les CR comprennent ces deux types de relations référentielles pourvu qu'elles réfèrent toutes au même référent, c'est-à-dire qu'elles entretiennent des relations coréférentielles. À partir de là, nous pouvons insister sur le fait que les CR sont fondées sur le principe de coréférence qui comprend des relations anaphoriques et des relations non anaphoriques. La citation de Charolles (1987 : 30) appuie cette proposition : « L'appartenance à une même chaîne suppose en effet, ainsi qu'on l'a déjà dit, qu'il y ait coréférence entre les expressions ».

Bilan

Ce chapitre initial avait pour objectif de fournir une base théorique pour notre étude, en définissant des notions connexes à celle de CR telles que *la cohérence*, *la cohésion*, *les plans d'organisation textuelle*, *la référence*, *le sens*, *la dénotation*, *la saillance*, *l'anaphore* et *la coréférence*.

Nous avons remarqué que, malgré le grand intérêt manifesté par les divers domaines d'étude, la définition de *la référence*, n'est pas consensuelle. Tenant compte de ce fait, nous

en reparlons succinctement afin de choisir notre orientation vers le sujet.

Après avoir examiné les deux conceptions sur la *référence*, la conception objectiviste et l'approche constructiviste (voir la section 2.2.1), et, en rappelant nos expériences de la vie réelle, nous décidons d'adopter un point de vue « ouvert », à l'instar d'auteurs comme Kleiber (1997), Larsson (1997) et Charolles (2002) : le langage offre la possibilité d'exprimer plusieurs types d'entités dans la mesure où l'homme peut fixer ces entités dans sa représentation mentale.

Ensuite, nous avons évoqué la notion de *saillance* et son rapport avec les CR. Considérant le fait que la saillance d'un référent peut influencer le choix de l'expression référentielle (voir les sections 2.2.2.3 et 2.3.2), il est tout à fait raisonnable de tenir compte de l'importance de la saillance dans le domaine des CR. L'auteur rédige son texte en choisissant attentivement des entités linguistiques pour qu'il soit lu et compris sans problème ; les lecteurs considèrent simultanément plusieurs facteurs pour trouver les bonnes relations coréférentielles existant dans le contexte et les comprendre sans équivoque. Dans ce sens, la représentation mentale de l'auteur et celle du lecteur interviennent d'une manière dynamique dans la constitution et l'interprétation des chaînes et le caractère multifactoriel de la saillance participe à ces processus.

En dernier lieu, nous avons distingué la notion de *CR* de celles d'*anaphore* et de *coréférence*. Pour construire une CR, le nombre minimum de maillons doit être de trois (voir la section 3.2.1 *supra*), sans quoi les notions d'*anaphore* et de *coréférence* suffisent à décrire les phénomènes (Schnecker et Landragin, 2014). Pourtant, quant au nombre de maillons, force est de reconnaître que la réponse absolument consensuelle n'est pas encore donnée, d'autant plus que la notion de *CR* est un sujet relativement récent dans le domaine de la linguistique référentielle. Nous suivons l'idée proposée par Schnecker (1997) et Schnecker et Landragin (2014) qui est la plus courante dans le domaine, afin de distinguer les simples paires de relation référentielle de la CR et de pouvoir fournir une analyse plus fine du phénomène.

Chapitre 2.

Caractéristiques des chaînes de référence

Introduction

Les notions de *CR* et de *maillon* étant déjà introduites dans la dernière section du chapitre précédent, nous proposons, dans ce présent chapitre, de démontrer les divers aspects du phénomène des CR.

Ce deuxième chapitre est divisé en quatre sous-sections. De prime abord, les caractéristiques des CR déduites par leur manifestation dans le texte seront détaillées (§1). Nous commencerons en abordant le problème de leurs bornes (§1.1) avant d'expliquer les diverses façons dont les différentes chaînes se déroulent et cohabitent au sein du texte (§1.2).

Le deuxième sous-thème a pour objectif de décrire les caractéristiques des maillons (§2). Pour ce faire, nous présenterons d'abord les différents points de vue sur les expressions référentielles susceptibles de s'employer comme maillon (§2.1), puis examinerons la corrélation entre le choix de maillons et les caractéristiques des CR (§2.2).

Ensuite, nous proposerons de mettre en évidence les principales conditions liées à la variété de matériau lexical des maillons (§3). Les trois facteurs, le genre textuel (§3.1), l'époque de rédaction (§3.2) et la langue de rédaction (§3.3), qui influencent le choix de matériau lexical des maillons, seront explicités par l'intermédiaire des études menées par nos prédécesseurs.

Nous finirons ce chapitre par une présentation de paramètres caractérisant les CR (§4). Les mesures inventoriées dans cette partie (§4.1) exploitent les données brutes fournies par le travail d'annotation et servent, de manière quantitative et qualitative, à analyser les CR (§4.2).

1 Les caractéristiques des chaînes

La notion de *CR* se distingue de celles d'*anaphore* et de *coréférence*, notamment par le nombre de maillons comme nous venons de le mentionner à la fin du chapitre précédent. Dans ce sens, la conception des *CR* concerne le « long terme référentiel » (Schneidecker et Landragin 2014 : 2).

Dans cette sous-section, d'un côté, le problème de la borne des *CR* sera discuté et d'un autre côté, nous présenterons cinq modalités de la coexistence des différentes chaînes dans le texte.

1.1 Les chaînes de référence et leurs bornes

Le fait que les *CR* possèdent au moins 3 maillons montre leur capacité à s'étendre tout au long du texte. Nous en proposons une illustration avec cet extrait, tiré du *Guide de tri* proposé par le gouvernement français¹. Il s'agit d'une explication brève de deux signalétiques présentées sur l'emballage :

(1) « Le Point Vert »

Contrairement à une idée reçue, **le Point Vert** ne veut pas dire que l'emballage est recyclable. **Il** signifie que *l'entreprise qui* met en vente ce produit participe financièrement à la collecte, au tri et au recyclage des emballages. **Le Point Vert** a été créé en France il y a plus de 20 ans dans le cadre de la Responsabilité Elargie du Producteur.

« L'info-tri Point Vert »

L'info-tri Point vert est la signalétique, proposée par Citeo, que *les entreprises* peuvent apposer sur *leurs* emballages afin de délivrer la bonne consigne de tri à *leurs* consommateurs. — (Quelle est la signalétique ?)

Le texte consiste en plusieurs rubriques. Dans le premier paragraphe qui concerne la signalétique « Le Point Vert », la *CR* « le Point Vert » est constituée par ses trois maillons signalés en gras dans l'exemple. Contrairement à cette chaîne qui ne dépasse pas son

¹ <https://www.consignesdetri.fr/baba/> (version 10.2019).

paragraphe thématique, la seconde chaîne de l'extrait, celle référant à « l'entreprise »², s'étend sur les deux paragraphes.

Dès lors, les questions suivantes semblent fondées : une chaîne persiste-t-elle tout au long d'un texte même dans le cas où celui-ci est un roman de plusieurs centaines de pages ? Dans le cas contraire, faut-il découper la chaîne lorsqu'un nouvel épisode ou un nouveau paragraphe débute ? Ces questions, déjà abordées dans les travaux de Schnedecker (1997) et de Schnedecker & Landragin (2014) entre autres, soulèvent un problème sur les processus cognitifs du lecteur : si on considère qu'une chaîne couvre tout un roman, le lecteur pourrait rencontrer le problème du coût du traitement que signale le commentaire suivant :

[...] dans quel état cognitif se trouverait le lecteur des trois tomes des *Frères Karamazov* qui aurait consciencieusement retenu ne serait-ce qu'une chaîne renvoyant à l'un desdits frères ? (Schnedecker, 1997 : 22-23)

Dans cette optique, le problème de la borne initiale ou finale des chaînes est évoqué dans le domaine de CR. De nombreuses formes de découpage d'un texte sont discutées (Capin et al., 2021 ; Schnedecker & Landragin, 2014 ; Schnedecker, 1997, 2021 ; entre autres), qu'elles soient typographiques, sémantiques, par épisodes ou par événements, etc., et le débat sur le sujet est loin d'être clos.

Ces questions suscitent deux types de réponses contradictoires comme citées par Schnedecker (1997 : 95) et Schnedecker & Landragin, (2014 : 6)³. D'un côté, les approches « collocatives » soulignent « une correspondance régulière entre les contextes informationnels et les types de stratégie référentielle (Marslen-Wilson, Levy et Komisarjevski, 1982 : 349, souligné par Schnedecker 1997 : 95), et d'un autre côté, les approches dites « déterministes » (*ibid.*) : elles mettent en valeur la structure textuelle qui pourrait conditionner le choix des expressions référentielles. Par exemple, d'après Ariel (1990), lorsqu'un texte est divisé en plusieurs paragraphes, les expressions référentielles de basse accessibilité telles que les noms propres ou les SN définis apparaissent à l'initiale des paragraphes : elles possèdent plus d'informations lexicales par rapport aux pronoms qui exigent plus de travail cognitif par manque de contenu descriptif (on aura l'occasion de le détailler dans la section

² Ici, nous ne distinguons pas le singulier et le pluriel parce que les deux sont toutes employées de manière générique : les maillons « l'entreprise » et « les entreprises » dans ce contexte sont interprétés comme le vendeur de produit mis dans l'emballage.

³ Cf. Capin et al. (2021).

2.1).

En outre, Schnedecker (2021 : 173-196) présente trois façons de marquer la délimitation des chaînes. En premier lieu, les différents plans de l'organisation textuelle peuvent provoquer une rupture dans la cohésion référentielle et donc avoir l'impact sur la composition des CR :

- découpage en paragraphe : comme montré dans l'exemple suivant que nous reprenons à Schnedecker (*ibid.*), le nom propre *l'Italie* se place au début de chaque paragraphe :

(2) **L'Italie**, en forme longue la République italienne (en italien *Italia* et *Repubblica Italiana*), est un pays d'Europe du Sud correspondant physiquement à une partie continentale [...].

L'Italie apporte une contribution très importante à la civilisation occidentale : elle est notamment le berceau de la civilisation étrusque, de la Grande-Grèce [...].

L'Italie, en ce début de XXI^e siècle, un acteur majeur de la scène internationale, forte de ses soixante millions d'habitants, de sa position de huitième puissance économique mondiale [...].

- distinction de la titrairie et du texte (dans certains genres textuels comme des textes de presse ou des textes scientifiques) : dans certains cas comme celui de l'exemple (3) ci-après, le titre d'un texte comprenant un pronom personnel ou un SN indéfini peut provoquer une rupture dans la CR :

(3) Titre : **Il** tente de s'immoler dans un terrain vague

Un jeune homme âgé de 29 ans, demeurant à Blainville-sur-l'Eau, en Meurthe-et-Moselle, a tenté de mettre fin à ses jours, samedi dernier, en s'immolant par le feu. (*Républicain lorrain*, in Schnedecker & Longo, 2012).

- découpage selon niveau énonciatif (discours rapportés), *etc.* : Schnedecker (2021 : 182) souligne que le discours rapporté « constitue une forme de découpage hybride » parce qu'il a une forme mixte des caractères typographique (il est marqué par les guillemets) et linguistique (il présente un caractère déictique et certaines marques d'oralité). Nous aurons l'occasion d'en reparler⁴.

⁴ Voir la note 22 *infra*.

Deuxièmement, l'auteur mentionne les découpages sémantiques effectués par les diverses expressions adverbiales cadratives. Celles-ci peuvent être temporelles (*en 1987, dans les années 90, actuellement,...*), spatiales (*à Strasbourg, dans le quartier de Robertsau, dans la communauté scientifique francophone,...*), et fictionnelles (*dans Harry Potter, dans la Casa de Papel,...*).

Finalement, elle souligne que les CR ont leurs propres bornes. En d'autres termes, l'apparition de certaines expressions référentielles comme le nom propre ou le SN démonstratif marquerait une rupture de chaîne et le redémarrage référentiel. D'après l'auteur (Schneidecker, 1997 : 73 et 2021 : 190), le nom propre est susceptible d'interrompre des chaînes par sa capacité à créer des ruptures discursives. Dans ces recherches, en analysant la relation entre le nom propre et la CR, une notion plus large, la *redénomination* est présentée comme la reprise des entités susceptibles de se manifester sous la forme des constructions comme « Npr... Npr... » et « SN... SN... ». Par la redénomination, le locuteur ouvre une nouvelle chaîne dans un contexte différent d'auparavant.

Ainsi, la borne des CR est un problème complexe qui concerne les divers éléments du texte et montre que le phénomène est étroitement lié non seulement à la structure du texte mais aussi aux caractéristiques des expressions référentielles employées.

1.2 Les chaînes de référence et leur déroulement dans le texte

Nous avons déjà constaté dans le chapitre précédent (voir la section 1.2 du chapitre 1, les plans d'organisation textuelle) que, dans un texte, différentes chaînes peuvent coexister. Autrement dit, elles peuvent se présenter ensemble de manières diversifiées. Nous nous inspirons des travaux de Schneidecker (1997 : 14 et 2021 : 100) et présentons cinq modalités de la coexistence des chaînes : succession, entrecroisement, dérivation, partition/fusion et parallélisme.

- Succession : Dans l'exemple (4), nous constatons la suite de plusieurs chaînes. La chaîne qui renvoie au personnage du jardinier débute juste après la fin de la première chaîne qui désigne le personnage du professeur. Puis, la troisième chaîne qui réfère au fils du jardinier commence à la fin de la deuxième chaîne :

(4) Johannesburg, années 70. **Un professeur d'Histoire** vit une vie de famille sans histoire, entouré de l'affection des **siens** et aveugle aux problèmes politiques et sociaux engendrés par l'« apartheid », régime ségrégationniste qui sévit en Union

sud-africaine. **Cet homme paisible a un jardinier noir**, paisible, **lui** aussi, et soumis à la fatalité. **Ce jardinier a un fils et ce fils** participe à une manifestation d'écoliers violemment réprimée par la police. On apprendra que **l'enfant** est arrêté, puis qu'**il** est mort. — (Résumé de « Une saison blanche et sèche », Télérama, 15/01/1992)⁵

- Entrecroisement : L'exemple (5) montre que les chaînes peuvent être instanciées consécutivement ou simultanément dans le texte. Les maillons de la chaîne qui renvoie à « Ursus » et ceux qui désignent « Homo » alternent au fur et à mesure du déroulement du texte :

(5) **Ursus** et *Homo* étaient liés d'une amitié étroite. **Ursus** était un homme, *Homo* était un loup. Leurs humeurs s'étaient convenues. **C'**était l'homme **qui** avait baptisé *le loup*. Probablement **il s'**était aussi choisi **lui-même son** nom ; ayant trouvé Ursus bon pour **lui**, **il** avait trouvé Homo bon pour *la bête*. — (extrait du roman « L'Homme qui rit » par Victor Hugo, tome VIII., p.7)⁶.

- Dérivation : La dérivation s'appuie sur une relation anaphorique. Le référent désigné par « les commerçants » dans l'exemple (6) *infra* entretient une relation associative avec « les chalets » ou bien « le marché de Noël » :

(6) *Le marché de Noël de Colmar* commencera une heure plus tard les jours de semaine pour l'édition 2022, et se terminera plus tôt. La décision a été prise en commun par les exposants et de l'office de tourisme. Cette année, les chauffages sont interdits dans *LES CHALETS* suite à la mise en œuvre d'une réglementation nationale. **Les commerçants** s'inquiètent, **ils** redoutent le froid, mais **ils** prennent acte de cette nouvelle règle. — (France Bleu, 21/09/2022)⁷

- Partition/fusion : Les différentes chaînes peuvent se diviser et fusionner au sein du texte. Dans (7), « les trois petits cochons » de la première chaîne est divisé en « le premier petit cochon », « le deuxième petit cochon » et « le troisième petit cochon » dans les chaînes qui suivent. Cependant, les deux premiers éléments divisés sont de nouveau regroupés

⁵ Cet exemple est emprunté à Schnedecker (1997 :13).

⁶ « L'Homme qui rit (éd. 1907) », *Wikisource*.

⁷ « Marchés de Noël et économies d'énergie : à Colmar, il commencera plus tard et fermera plus tôt le soir » <https://www.francebleu.fr/infos/environnement/marche-de-noel-et-economie-d-energie-a-colmar-il-commencera-plus-tard-et-fermera-plus-tot-1663771991>

dans le paragraphe suivant par une expression « des deux premiers petits cochons ».

- (7) **Les trois petits cochons** veulent vivre **leur** vie et quittent le foyer familial pour tenter **leur** chance dans le monde. **Le premier petit cochon** se construit une maison de paille. **Le deuxième petit cochon** se construit une maison faite de bois. **Le troisième petit cochon** se construit une maison de briques et de ciment.

LE GRAND MÉCHANT LOUP parvient à détruire les maisons **des deux premiers petits cochons** en soufflant dessus et **les** dévore. En revanche, IL est impuissant contre celle **du troisième petit cochon**. — (extrait de « Les Trois Petits Cochons »)⁸

• Parallélisme : Lorsque deux référents sont « traités sur un pied d'égalité et présentés comme aussi importants l'un que l'autre » (Schneidecker, 2021 : 101), il s'agit des chaînes parallèles. Elles peuvent être représentées dans le texte par le biais des pronoms corrélés « l'un... l'autre » ou « le premier... le second » (Schneidecker, 2006) ou par la même fonction syntaxique que chaque maillon exerce dans la phrase. Regardons les deux exemples suivants que nous reprenons à Schneidecker (2021 : 102) :

- (8) Autrefois, dans les spéculations sur les terrains de la plaine Monceau, Saccard avait eu des discussions, toute une brouille même avec Gundermann. ILS ne pouvaient s'entendre. **L'un**, passionné et jouisseur, *l'autre*, sobre et d'une froide logique. Aussi **le premier**, dans sa colère, exaspéré encore par cette entrée triomphale, s'en allait-il lorsque *l'autre* l'appela. — (extrait de « L'argent », Zola)⁹

- (9) Si la misère humaine pouvait être résumée, elle l'eût été par GWYNPLAINE ET DEA. ILS semblaient être nés chacun dans un compartiment du sépulcre ; **Gwynplaine** dans l'horrible, *Dea* dans le noir. LEURS existences étaient faites avec des ténèbres d'espèce différente, prises dans les deux côtés formidables de la nuit. Ces ténèbres, *Dea* les avait en *elle* et **Gwynplaine** les avait sur **lui**. Il y avait du fantôme dans *Dea* et du spectre dans **Gwynplaine**. *Dea* était dans le lugubre, et **Gwynplaine** dans le pire. ... — (extrait de « L'homme qui rit », II, 4., Hugo)¹⁰

⁸ https://fr.wikipedia.org/wiki/Les_Trois_Petits_Cochons.

⁹ Nous reprenons une partie de l'exemple de Schneidecker (2021 : 102).

¹⁰ Schneidecker (2021 : 102).

Schnedecker (2021 : 102), en citant son travail antérieur (Schnedecker, 2006)¹¹, souligne que ces différentes modalités de coexistence des CR dans un texte peuvent servir de paramètres permettant de « caractériser des genres ou des référents, des époques ou des systèmes linguistiques » particuliers. En partant de cette idée, nous proposerons, tout au long du présent chapitre, les éléments susceptibles de montrer la corrélation entre les caractéristiques du texte et celles des CR.

2 Les caractéristiques des maillons

Les maillons sont des expressions référentielles qui constituent une CR. Autrement dit, ce sont des unités linguistiques qui construisent les CR. Quelles sont les caractéristiques de ces éléments linguistiques ? Toute expression référentielle peut-elle fonctionner comme un maillon ? Sinon, y a-t-il des formes typiques des expressions référentielles qui peuvent être utilisées comme des maillons ?

Avant d'entamer une discussion autour de ces questions nous présenterons, à grands traits, différents points de vue sur les expressions référentielles, ceux de Charolles (2002) et d'Ariel (1990), entre autres. En outre, en renvoyant à Schnedecker (2005), nous expliciterons deux sortes de CR distinguées en fonction de la diversité des types d'expressions référentielles utilisées comme maillons. L'ensemble de ces descriptions fournira une base de compréhension sur le statut et le rôle des expressions référentielles dans la constitution des CR.

2.1 Expressions référentielles susceptibles de servir de maillon

Nous rappelons brièvement notre position sur les conceptions de *référence* et d'*existence* pour aboutir à une discussion de la nature des maillons : pour qu'une expression

¹¹ L'auteur montre que la mise en chaîne des expressions référentielles peut varier selon le genre textuel : les constructions parallèles en alternance sont observées dans les textes narratifs pour un effet de contraste tandis que dans les textes informatifs-argumentatifs, les chaînes de référence des éléments sériés sont souvent présentées « en bloc ».

réfère, il faut qu'elle soit associée à une entité à même d'être fixée dans la représentation mentale.

Sous cet angle, Charolles (1988) insiste sur le fait que seules les expressions nominales (ou pronominales) susceptibles d'identifier une quelconque entité d'un texte, « quelle que soit sa forme d'existence (personne humaine, événement, entité abstraite ...) », peuvent faire partie d'une chaîne (*ibid* : 8). Puis, dans un ouvrage ultérieur (2002), il classe cinq catégories d'expressions référentielles : les noms propres, les SN définis, démonstratifs, indéfinis et les pronoms. Son optique sur la nature des maillons est relativement stricte par rapport à ce que nous introduirons prochainement.

La figure 4 *infra* décrit la théorie de l'Accessibilité d'Ariel (1990). Cette théorie offre des possibilités plus étendues par rapport au classement de la catégorie grammaticale d'expressions référentielle de Charolles (2002) : elle permet d'avoir une vue plus ouverte sur la limite ou la variété des expressions référentielles.

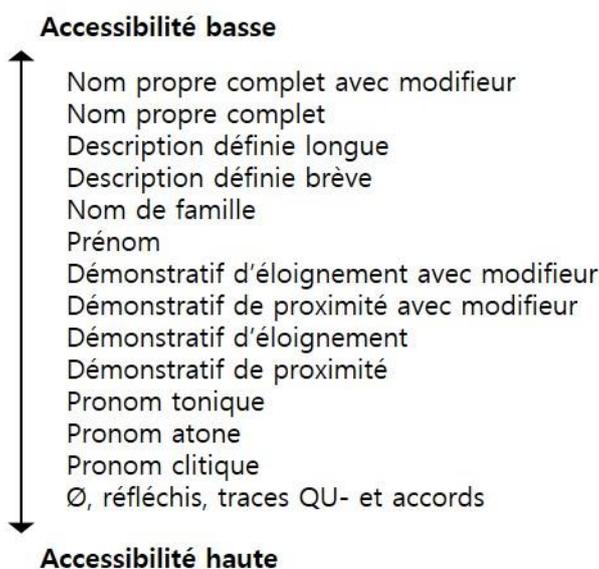


Figure 4 Echelle de l'accessibilité (Ariel, 1990)

Comme nous le voyons dans cette figure sur l'accessibilité des expressions référentielles, selon Ariel (1990), les éléments linguistiques tels que l'anaphore zéro, les phénomènes d'accord et même les prédicats verbaux sont tous considérés comme expressions référentielles.

À part ces deux approches, Schnedecker (2021 : 123-125) présente une conception encore plus ouverte que celle d'Ariel, qui comprend toutes les formes de rappels comme les

appositions et les phénomènes de coréférence réalisés *via* des participes présents ou passés, ainsi qu'une autre conception qui se situe dans la zone médiane, celle proposée par Landragin (2011, cité dans Schnedecker, 2021 : 125). Dans cette dernière position, les phénomènes d'accord verbaux et participiaux sont exclus, et l'auteur introduit une nouvelle notion, celle de « maillon faible », qui désigne les maillons sans forme linguistique.

La délimitation des expressions référentielles, en d'autres mots, définir la nature des maillons, est une étape essentielle dans les recherches portant sur les CR. En fonction du choix de l'auteur, les composants d'une chaîne se manifesteront différemment. Notre perspective sur la nature des maillons sera précisée dans le chapitre 4, où nous abordons les expressions référentielles, ainsi que dans le chapitre 6, dans lequel nous discuterons des conventions d'annotation.

2.2 Diversité des types de maillons et caractéristiques des chaînes

Les CR peuvent être distinguées par leur matériel de base, à savoir par les types de maillons employés. Les chaînes dites homogènes sont composées massivement de noms propres réitérés et de pronoms, tandis que les chaînes hétérogènes sont constituées de pronoms, des syntagmes nominaux à tête lexicale (le nom propre, SN défini, et SN démonstratif) et d'un matériau lexical diversifié (Schnedecker, 2005 : 96-97). La chaîne de « Ramon » de l'exemple (10) est homogène du point de vue des contenus :

(10) Malgré le sentiment de compassion qu'**il** avait ressenti à la fin de leur rencontre dans le jardin du Luxembourg, **Ramon** ne pouvait rien changer au fait que D'Ardelo appartenait à la sorte de gens qu'**il** n'aimait pas. Et cela, même s'ils avaient tous les deux quelque chose en commun : la passion d'éblouir les autres ; de les surprendre par une réflexion amusante ; de conquérir une femme sous leurs yeux. Sauf que **Ramon** n'était pas un Narcisse. — (Milan Kundera, *La fête de l'insignifiance*, 2013, pp.68-69)

Dans cet exemple, le pronom personnel « il » et le nom propre « Ramon » sont utilisés comme expressions référentielles renvoyant à « Ramon ». Regardons l'exemple (11) qui montre comment différentes catégories grammaticales peuvent servir de maillons dans une chaîne hétérogène :

(11) Derrière **ses** lunettes cerclées, et **ses** faux airs de Harry Potter asiatique, **Moon Jae-in** cache une détermination en acier trempé dans les drames de la division de la

péninsule coréenne. Depuis un an, **cet ancien soldat des forces spéciales sud-coréennes** joue avec dextérité l'entremetteur entre les deux plus imprévisibles fauves de la scène internationale : Donald Trump et Kim Jong-un. Un travail inlassable qui se poursuit à Paris, où **le président sud-coréen** entame lundi à l'Élysée, **sa** première tournée européenne, nouvelle étape de **son** offensive diplomatique tous azimuts en faveur de la réconciliation avec Pyongyang. — (*Le Figaro*, 14/10/2018)¹²

Le référent « Moon Jae-in » de l'exemple (11) est repris par le syntagme démonstratif « cet ancien soldat des forces spéciales sud-coréennes », par les déterminants possessifs et par le syntagme nominal défini « le président sud-coréen » dans la phrase qui suit. Dans les chaînes hétérogènes, l'emploi des expressions référentielles est diversifié, et les syntagmes nominaux employés contiennent souvent des informations descriptives sur le référent.

Ainsi, le type et la substance des maillons ont des répercussions sur les caractéristiques des chaînes, voire sur celles du texte. Nous examinerons en détails la corrélation entre le maillon, la chaîne et le texte dans le sous-thème suivant.

3 Trois facteurs relatifs à la variété du matériau lexical

Le matériau lexical des maillons peut être très varié comme l'illustre l'exemple (11) que nous venons de voir. D'après Schnedecker & Landragin (2014 : 8-11), la variété de matériau lexical dépend de trois facteurs : le genre textuel (Schnedecker, 2005, 2014, 2021 ; Longo & Todirascu, 2014 ; Schnedecker & Longo, 2012), la date de la composition (Combettes, 2006 ; Schnedecker, 2021 ; Capin *et al.*, 2021 ; Glikman *et al.*, 2014 ; Oberlé *et al.*, 2018) et la langue de rédaction du texte (Baumer, 2015 ; Lundquist, 2005 ; Schnedecker, 2021 ; Baumer *et al.*, 2020). Inspirée par cette approche, nous parlerons, dans les points suivants, de ces trois facteurs et de leur manière d'affecter les CR.

¹² « Moon Jae-in, l'apôtre tenace d'une réconciliation entre les deux Corées ».

<http://www.lefigaro.fr/international/2018/10/14/01003-20181014ARTFIG00133-moon-jae-in-l-apotre-tenace-de-la-reconciliation.php>

3.1 Le genre textuel

Le genre textuel conditionne divers paramètres des CR, notamment « l'apparition et la fréquence de certaines marques linguistiques » (Schneidecker, 2021 : 163). Les paragraphes suivants récapitulent quelques travaux du domaine qui portent soit sur un genre particulier, soit sur la comparaison inter-genres.

En premier lieu, les études sur un genre particulier choisissent plusieurs textes du même genre textuel, les analysent et dégagent la tendance générale de ce genre sous l'angle des CR. Ainsi, plusieurs genres se distinguent : les discours instructionnels (Maes, Arts & Noordman, 2004), les documents techniques (Dupont & Bestgen 2006), le portrait journalistique (Schneidecker, 2005), le fait divers (Schneidecker et Longo, 2012)¹³, les textes administratifs et juridiques (Longo et Todirascu, 2014), les nouvelles contemporaines (Baumer, 2017) et les romans réalistes du 19e siècle (Baumer, Dias & Schneidecker, 2020).

En plus de ces approches, la comparaison inter-genres suscite l'intérêt des chercheurs. Schneidecker (2021 : 163) présente l'une des approches récentes en la matière (cf. Kunz, 2010) menée sur des textes de fiction, de vulgarisation scientifique et des essais en anglais et en allemand. D'après cette recherche, les textes de fiction montrent « une continuité référentielle très dense » par la longueur et la composition des chaînes par rapport aux autres types de textes. Cela peut être explicité par le fait que les genres narratifs (le fait divers, la nouvelle, le roman) possèdent un ou des personnage(s) dominant(s) en tant que référent central. La récurrence de ce référent central couvre tout le texte et il est plus saillant que d'autres référents, ce qui facilite « le suivi référentiel » (Schneidecker, 2014 : 24).

Tutin (2002) s'intéresse à l'apparition et la résolution des expressions anaphoriques dans quatre domaines de discours¹⁴ différents : 4 articles d'un journal en sciences humaines, 10 articles sur l'économie du *Monde*, un manuel technique d'imprimante et un extrait de roman. Le tableau suivant est une partie de l'analyse qui décrit la distribution des expressions référentielles dans les textes de différents genres.

¹³ Ces articles qui viennent d'être mentionnés sont cités dans Schneidecker & Landragin (2014 : 8-9). Nous ajouterons à cette liste quelques autres recherches relatives à ce domaine.

¹⁴ Tutin (2002) parle du « genre textuel » dans son travail mais, à proprement parler, les articles de journaux ne sont pas un genre textuel. Il en va de même pour le manuel technique. Nous utilisons donc le terme « domaine de discours » pour désigner les quatre registres textuels de son corpus.

	Human science papers Hermès	News-paper Le Monde Eco	Technical manual HP Printer manual	Literary Text J. Vernes' text
“Clitic” p. pronouns	46	52	68,5	54
Possessive determiners	26	31	4	36,5
Demonstrative pronouns	10	6	18,5	3
Disjunctive p. pronouns	6	4	0	3
Noun head ellipses	7	3	4	2
Numeral pronouns	0,5	2	0	2
Indefinite pronouns	3	1,5	5,5	2

Tableau 2 Distribution des expressions anaphoriques (Tutin, 2002)

Une autre comparaison inter-genres (domaines de discours, à proprement parler) a été réalisée par Longo et Todirascu (2010) sur des articles du *Monde*, des éditoriaux du *Monde Diplomatique*, un roman, des textes administratifs, ainsi que sur quelques textes à caractère législatif de l'*Acquis Communautaire* et des rapports publics. Les auteurs analysent les cinq domaines de discours variés selon cinq critères, ainsi que la longueur moyenne des chaînes (le nombre des expressions référentielles renvoyant à une même entité), la distance moyenne entre les éléments de chaîne (le nombre de phrases qui séparent les maillons), la fréquence des mentions selon leur catégorie grammaticale, la catégorie grammaticale du premier maillon d'une chaîne et l'identité entre le thème de la phrase et de premier maillon d'une chaîne. Un des résultats remarquables est que les chaînes dans le roman sont en moyenne trois fois plus longues que celles dans les textes informatifs comme, par exemple, le texte législatif. Ce résultat, d'après les auteurs (*ibid* : 61), peut être justifié par le fait que le texte législatif, en règle générale, comprend plusieurs référents concurrents : il y a plusieurs sous-thèmes et plusieurs actants dans le texte, et par conséquent, un bon nombre de chaînes

débutent et disparaissent au fur et à mesure du texte¹⁵. En revanche, dans les romans, tout au long du texte, l'histoire se déroule autour des personnages principaux et corollairement, les chaînes sont moins nombreuses mais relativement plus longues que celles des textes législatifs¹⁶.

Schnedecker (2014) présente également une analyse des modalités des expressions référentielles d'un point de vue différent. Considérant le fait que la plupart des études inter-langues s'appuient sur les genres à dominante narrative, et que souvent elles analysent la relation entre genre et CR de manière trop globale, l'auteur choisit deux genres textuels non-narratifs : les introductions de textes encyclopédiques et les recettes de cuisine. Ces genres ont des objectifs et des structures différents : l'un incite à l'action et la dirige pas à pas, tandis que l'autre vise à transmettre des connaissances, des aspects divers sur un thème au public. En conséquence, cette analyse insiste sur le fait que les CR dans les recettes de cuisine sont en général « brèves avec des maillons en relation d'étroite proximité, mais elles restent limitées à un paragraphe (bref lui aussi) » (Schnedecker, 2014 : 39). À l'opposé, les CR dans les introductions de textes encyclopédiques couvrent l'ensemble du texte et leurs maillons sont plus espacés.

Les études précitées montrent que les CR sont tributaires des genres textuels dans lesquels elles se manifestent. Regardons les exemples suivants afin d'illustrer le phénomène. Dans les portraits journalistiques, comme dans (12), les noms propres, les pronoms et les syntagmes nominaux sont majoritairement employés :

(12) Dix ans, 1 000 toiles, 2 000 dessins... En une décennie, **Jean-Michel Basquiat** a bouleversé la peinture contemporaine. Une comète venue de nulle part ? Plutôt le fruit d'une époque, et d'une immense culture. **Jean**, comme on l'appelait simplement, est né en 1960 d'un père haïtien et d'une mère d'origine portoricaine, dans la petite bourgeoisie de Brooklyn.

Il grandit auprès des artistes de la rue, mais aussi des grands noms de la peinture. Du MoMA au Metropolitan en passant par le Brooklyn Museum, **il** écume les musées, encouragé par une mère fière de **ce gamin** « si éveillé, si intelligent, un esprit absolument extraordinaire... **Il** a dessiné et **peint** toute **sa** vie dès l'âge de 3 ou 4 ans

¹⁵ À ce propos, les auteurs mentionnent brièvement les points importants : elles n'ont malheureusement pas mis l'exemple dans leur article.

¹⁶ *Idem*.

», se souvenait-elle. — (*Le Monde*, 02/10/2018)¹⁷

Ensuite, suivant Longo & Todirascu (2014), les textes administratifs et juridiques, très attentifs à éviter les ambiguïtés, utilisent fréquemment « des procédés de répétition lexicale de SN définis » (*ibid* : 96) et « pas ou peu de reprises pronominales » (*ibid* : 88). L'extrait ci-dessous montre indéniablement cette caractéristique du genre :

(13) **Le changement climatique** aura des répercussions économiques et sociales importantes, notamment dans certaines régions et certains secteurs qui seront touchés plus sévèrement que d'autres. Une partie de la population (les personnes âgées, les handicapés, les ménages à faible revenu) devrait aussi souffrir davantage.

La lutte contre le changement climatique doit se faire à deux niveaux. Il s'agit d'abord et avant tout de réduire les émissions de gaz à effet de serre (au moyen de mesures d'atténuation), puis de prendre les mesures d'adaptation qui s'imposent pour faire face aux conséquences inévitables de **ce changement**. (Corpus « Rapport public » - cité par Longo & Todirascu, 2014 : 84)

À part les aspects lexicaux, plusieurs d'autres facteurs peuvent être utilisés comme critères de comparaison (nous aurons l'occasion d'en reparler à la fin de ce chapitre) car chaque genre textuel a sa propre manière de présenter les CR.

3.2 L'époque de rédaction

L'époque de rédaction des textes concerne également la diversité du matériau lexical des CR. Perret (2000 : 17) signale que l'ancien français privilégie la stabilité désignationnelle. Plus spécifiquement, le matériau lexical des expressions référentielles constituant une chaîne est moins diversifié en ancien français qu'en français moderne. L'auteur introduit le concept de « coefficient de stabilité » pour rendre compte de ce phénomène. Cette notion concerne la diversité des expressions référentielles employées afin de désigner un référent :

Obtenu en divisant, pour un référent donné (un personnage), le nombre total de l'anaphores nominales par le nombre de désignations différentes... Plus le coefficient

¹⁷ « Jean-Michel Basquiat, l'effervescence d'une comète »

https://www.lemonde.fr/arts/article/2018/10/02/jean-michel-basquiat-l-effervescence-d-une-comete_5363368_1655012.html

de stabilité est élevé, moins il y a de désignations différentes par rapport au nombre d'anaphores et donc plus la stabilité référentielle est grande. (Perret 2000 : 17)

Comme le signale Perret, le chiffre élevé signifie que les expressions référentielles d'une chaîne sont plutôt monotones. En ancien français, les mêmes types d'expressions référentielles sont employés pour que la procédure de la référence soit plus simple et stable. De ce fait, dans les textes écrits aux 14^{ème} et 15^{ème} siècles, il y a moins de richesse dans les désignations que dans le texte de Proust étudié par l'auteur et les reprises nominales répétitives se présentent largement au fil du texte.

À partir de cela, plusieurs recherches ont démontré que les CR se présentent différemment selon l'époque où le texte a été écrit. Les recherches visant à montrer ces aspects des CR travaillent sur le corpus d'un même genre textuel. Obry *et al.* (2017 : 108) analysent les CR dans des récits brefs anciens (13^{ème}, 15^{ème}, et 16^{ème} siècles) à tonalité comique : considérant le genre textuel du corpus, le travail d'annotation de cette étude est fait sur toutes les mentions d'animés humains. Cette recherche montre qu'il existe « une évolution diachronique dans la composition des CR ». Par exemple, l'emploi d'un SN indéfini en tant que premier maillon est réservé aux chaînes des personnages centraux dans les textes du 13^{ème} siècle parce que, dans les textes de cette époque, l'article indéfini « semble conférer une saillance au référent » (*ibid* : 108). Pourtant, au fil du temps, ce rôle de l'article indéfini commence à s'affaiblir et le nom propre « paraît le remplacer dans ce rôle » (*ibid* : 108). À partir du 15^{ème} siècle, Apparaissent également de nouvelles expressions référentielles telles que *ledit* et *lequel*. L'exemple suivant est une partie du texte du 15^{ème} siècle présenté dans le travail de Obry *et al.* (2017 : 107) :

(14) En la ville de Valenciennes eut nagueres **ung notable bourgeois**, en **son** temps receveur de Haynau, **lequel** entre les aultres fut renommé de large et discrete prudence. Et entre **ses** loables vertuz celle de liberalité ne fut par la maindre, car par icelle **vint** en la grace des princes, seigneurs et aultres gens de tous estaz. (Corpus « Cent nouvelles nouvelles », composées vers 1460)¹⁸

En outre, le travail d'Oberlé *et al.* (2018) examine l'évolution diachronique des CR dans un autre genre textuel. À travers un corpus constitué de 20 textes encyclopédiques diachroniques (du 12^{ème} siècle jusqu'au 20^{ème} siècle), les auteurs analysent les caractéristiques des textes de chaque époque et observent la corrélation entre les CR et la

¹⁸ Repris de Obry *et al.* (2017 : 107).

structure textuelle. Selon eux, l'évolution du système linguistique et celle du genre sont des facteurs qui influencent les caractéristiques des chaînes. Par exemple, en règle générale, le pronom personnel est une catégorie grammaticale dominante parmi tous les pronoms employés dans les textes encyclopédiques, à l'exclusion des textes des 14^{ème} et 15^{ème} siècles. Pendant cette période, le pronom zéro est employé à peu près deux fois plus que le pronom personnel, ce qui peut être expliqué par la particularité du français de cette époque qui n'exprime pas forcément le pronom personnel sujet. Ensuite, la diminution de l'emploi du pronom personnel est observée entre les 18^{ème} et 20^{ème} siècles : la proportion atteint 40 % au 18^{ème} siècle et arrive à 17 % au 20^{ème} siècle. Afin de comprendre cette disparité, les auteurs parlent des traits du genre qui diffèrent selon l'époque. Les textes encyclopédiques du 18^{ème} siècle ont des séquences narratives longues, et par conséquent, le référent principal est souvent repris par le pronom personnel, tandis que les textes du 20^{ème} siècle, notamment ceux de *Wikipedia*, qui est un travail collectif, se composent souvent de paragraphes courts, rédigés par plusieurs personnes, « dont les SN sont dénués de reprise pronominale » (Oberlé *et al.*, 2018 : 99).

Le travail de Capin *et al.* (2021) étaye cette perspective diachronique. Les auteurs examinent le rôle de la composition des CR dans la segmentation et la structuration des textes à l'aide de sept textes à dominante narrative de la période du 12^{ème} au 20^{ème} siècle. Les textes anciens n'ayant pas de paragraphes au sens moderne montrent une nette différence par rapport aux textes modernes au niveau de la modalité selon laquelle les CR participent à la structuration du texte. Dans les textes anciens, l'occurrence d'un Npr à la tête d'une CR indique la délimitation des blocs thématiques, tandis que, dans les romans du 19^{ème} et du 20^{ème} siècles qui sont des textes polytypologiques, la redénomination par le Npr ou le SN ne montre pas uniquement le changement de structure textuelle, mais concerne également le changement de focalisation.

Ainsi, comme le système linguistique, la façon de structurer un genre textuel, la segmentation du texte, la mode de production et la visée du texte changent au fil des siècles : les CR se présentent différemment selon l'époque.

3.3 La langue de rédaction

La diversité du matériau lexical des maillons diffère selon la langue de rédaction du texte. Certaines études comparent les emplois des expressions anaphoriques dans des textes de

langue romane vs germanique.

Dans la recherche comparative de Lundquist (2005) menée sur le français et le danois, l'auteur insiste sur le fait que les normes rhétoriques et stylistiques qui sont étroitement liées à une culture donnée sont l'un des éléments importants dans la textualisation (*ibid* : 73). D'après cette étude, les anaphores dites fidèles¹⁹ sont plus utilisées en danois qu'en français. Les normes du français préfèrent varier les renvois anaphoriques dans un texte pour une raison esthétique et divertissante, tandis que le danois envisage davantage la clarté et l'uniformité des expressions anaphoriques en employant souvent les expressions référentielles composées du même matériau lexical que leur antécédent. L'auteur choisit quatre textes produits dans des conditions identiques dont deux français et deux danois. Il s'agit de textes traitant d'un même sujet (corpus dit comparable), la situation politique en Argentine autour de l'élection présidentielle dont la chaîne principale concerne l'un des deux candidats présidentiels. Les résultats montrent que, dans les textes danois, seules 2,5% des 80 anaphores sont des anaphores infidèles alors que, dans les textes français, les anaphores infidèles sont approximativement six fois plus employées. L'une des différences plus générales et typologiques entre les deux langues est que le français a tendance à faire varier les SN et à faire l'économie des verbes alors que le danois préfère des continuations fidèles, à savoir des SN avec moins de contenus sémantiques variés.

Dans cette optique, Baumer (2015) réalise une étude contrastive des points de vue inter-genres et inter-langues : il aborde les CR dans les nouvelles et portraits journalistiques, dans le cadre d'une comparaison français/anglais. L'auteur analyse les expressions référentielles employées renvoyant au personnage principal²⁰ de chaque texte. En gros, la distribution des expressions référentielles est identique pour chaque genre de chaque langue. Le taux d'emploi des expressions référentielles s'ordonne comme suit : « Pronom > Nom propre > Anaphore nominale » dans les nouvelles et portraits journalistiques des deux langues. Globalement, les textes français, par rapport aux textes anglais, tendent à utiliser plus fréquemment les formes dites « pleines » que sont les noms propres et les anaphores

¹⁹ Apothéloz (1995 : 36) définit l'anaphore fidèle pour des cas où « un référent préalablement introduit dans le texte est rappelé au moyen d'un SN défini ou démonstratif dont le nom tête est celui-là même au moyen duquel il a été introduit ». ex. *Pierre a mangé une pomme ce matin. Cette pomme était celle qu'il avait gardé dans son sac depuis quelques jours.*

²⁰ Comme nous l'avons déjà mentionné plus haut (voir la section 3.1) en citant Schnedecker (2014 : 24), dans des genres à dominante narrative, le personnage principal fait office de référent central.

nominales. Comme le montre le tableau suivant (tableau 3), dans les textes anglais, le taux d'emploi des pronoms est plus élevé que dans les textes français et il n'y a pas de grande différence inter-genres à ce niveau-là (Baumer, 2015 : 43-44 et 162-163). Au contraire, les textes français montrent un peu plus de décalage inter-genres dans la distribution des expressions référentielles.

	Textes anglais			Textes français		
Nouvelles	Pro	Npr	AN	Pro	Npr	AN
	70,8% > 24,4% > 4,8%			64,2% > 26,4% > 9,4%		
Portraits journalistiques	Pro	Npr	AN	Pro	Npr	AN
	70,5% > 26,1% > 3,5%			59% > 22% > 19%		

Tableau 3 Récapitulatif de la distribution des ER analysée par Baumer (2015, 43-44 & 162-163)

Un autre point plus concret souligné dans cette recherche, afin d'étayer l'hypothèse de Lundquist, est que, dans les portraits journalistiques, surtout quand l'auteur exprime une « subjectivité indirecte », c'est-à-dire l'intervention subjective de l'énonciateur sur le référent, la proportion du nom propre en anglais est supérieure à celle en français. En anglais, la subjectivité de l'énonciateur est exprimée soit par les énoncés qui comportent des jugements appréciatifs, soit par l'emploi de la construction passive, alors qu'en français, ce sont les anaphores intégrées dans les énonciations qui jouent ce même rôle. Ce point est facilement observé lorsqu'on compare deux portraits journalistiques décrivant le même personnage, comme dans les deux exemples qui suivent :

(15) **Tomi Ungerer**, the award-winning French artist and satirist, published more than 140 books in German, French and English, criss-crossing genres.

His oeuvre included anti-Vietnam War posters, darkly comic children's books, a mischievous rethinking of *The Joy of Sex*, children's literature and a 1986 collection titled *Guardian Angels of Hell*, featuring interviews with sex workers at a Hamburg brothel.

Ungerer was also a sculptor, printmaker, painter, caricaturist and antique toy collector, an ad man for the Ice Capades (travelling ice-skating shows that started in the US in the Forties and died out in the Nineties), and a food editor for *Playboy* magazine.

In the German city of Karlsruhe, **he** designed a kindergarten in the shape of a cat, with a whiskered nose, windows for eyes, a door for a mouth and a slide for a tail.

Raised in the Alsace region, **Ungerer** lived under Nazi occupation and hitchhiked

across Europe before moving to New York in 1956. **He** found work as an illustrator for publications including *The New York Times*, *Life* and *Harper's Bazaar*, while creating posters for films such as Stanley Kubrick's *Dr. Strangelove*.

Ungerer became best known for **his** children's books, including the illustrations for Jeff Brown's 1964 classic *Flat Stanley*, about a boy who is crushed flat by a bulletin board, slips inside envelopes to travel by mail, and restores himself to proper size with the aid of a bicycle pump. — (*Independent*, 18/02/2019)²¹

(16) **Il** ne comptait pas arrêter de dessiner de sitôt. Il y a encore quelques semaines, **le dessinateur et illustrateur alsacien Tomi Ungerer** dévoilait de nombreux projets au « Dernières Nouvelles d'Alsace ». Mais **l'auteur de célèbres ouvrages pour enfants** tels que « Les trois brigands » (1961) ou « Jean de la Lune » (1966), est décédé à l'âge de 87 ans dans la nuit de vendredi à samedi, à Cork, en Irlande, au domicile de **sa** fille Aria. **Il** a été retrouvé le matin, à 9 heures, dans **son** lit, selon les « Dernières nouvelles d'Alsace ». ...

Né à Strasbourg en 1931, **il** a 8 ans lorsque la guerre éclate. On imagine un traumatisme. C'est mal le connaître : « **Je** comparerais plutôt **mon** enfance à un safari : j'ai vécu dans la poche de Colmar une bataille de trois mois, des explosions, des bombardements, c'était passionnant pour un gamin ! »²² Après la guerre, au moment

²¹ « Tomi Ungerer: Illustrator and author behind legendary posters and children's books »

<https://www.independent.co.uk/news/obituaries/tomi-ungerer-dead-obituary-illustrator-author-childrens-literature-political-erotic-kubrick-the-three-a8777226.html>

²² Le traitement des expressions référentielles, notamment des pronoms de première et deuxième personnes, dans des discours rapportés tels que dialogues romanesques, des interviews, entre autres, soulève une problématique : devons-nous considérer les mentions des discours représentés comme participants d'une même CR ? (Schneidecker, 2021 : 79).

Dans la section 1.1 de ce chapitre, nous avons brièvement mentionné le fait que le discours rapporté peut être un des moyens de découper la chaîne. Schneidecker (*ibid* : 182) souligne que le discours rapporté forme, en tant que tel, un « 'bloc fortement délimité, au sein d'un discours enchâssant ». Au vu son statut particulier dans le texte, la question se pose sur le traitement des maillons présentés dans le discours rapporté : soit de traiter toutes les expressions coréférentielles comme maillons d'une même chaîne soit distinguer les maillons selon « leur niveau énonciatif » (*ibid* :183).

Nous considérons que les expressions référentielles dans le discours rapporté ont une relation coréférentielle non seulement entre eux, mais aussi avec celles qui sont en dehors de ce discours malgré le changement énonciatif : dans cet exemple précis, « je » et « mon » dans le discours rapporté de « Tomi Ungerer » désignant lui-même sont

où les Alsaciens sont traités de « sales Boches », **Ungerer** ne le supporte pas.

A 25 ans, **il** s'échappe à New York avec **ses** cartons à dessin. **Il** a hérité de **son** père, un ingénieur en horlogerie, le goût des arts et la passion du crayon. Nous sommes alors en 1956, période glorieuse pour la presse, l'édition, et les illustrateurs : en quelques mois, **le jeune Tomi s'**impose. **Il** dessine pour les magazines, puis **se lance** dans les livres pour enfants. Chacun d'eux est un succès public. Mais **l'homme** ne fait pas que ça. Artiste engagé, **il** signe des affiches politiques — contre la ségrégation ou la guerre du Viêt Nam notamment — très virulentes.

Surtout, **ce libertin** « fait les 400 coups » et les **dessine** également. A l'époque, Internet n'existe pas et, longtemps, personne ne fera le rapprochement. Jusqu'à la publication, en 1969, du « Fornicon », recueil de dessins érotiques. – (*Le Parisien*, 09/02/2019)²³

Ces deux exemples sur la personnalité « Tomi Ungerer » montrent que l'anaphore infidèle est plus souvent utilisée en français qu'en anglais. Le journaliste français fournit l'information supplémentaire qu'il veut diffuser sur le référent sous la forme d'anaphore nominale (par les SN) ; le journaliste anglais, au lieu d'employer des anaphores nominales, tente d'illustrer la situation par une description verbale.

Comme décrit tout au long de ce sous-thème, la variété des matériaux lexicaux des CR dépend de trois facteurs qui sont le genre textuel, la date de rédaction et la langue de rédaction. L'objectif de cette étude pourra être explicité par ce fait : nous analyserons les CR présentées dans les romans français et coréen du début 20^{ème} siècle (nous aurons l'occasion d'en reparler plus en détail dans le chapitre 5).

coréférentiels avec d'autres maillons qui renvoient au même personnage à l'extérieur de ce discours.

²³ « Décès de Tomi Ungerer, auteur et dessinateur de livres jeunesse »

<http://www.leparisien.fr/culture-loisirs/livres/les-trois-brigands-jean-de-la-lune-l-auteur-et-dessinateur-tomi-ungerer-est-mort-09-02-2019-8007987.php>

4 Un ensemble de paramètres pour caractériser les chaînes de référence

Après avoir présenté brièvement les facteurs qui affectent les CR et leurs maillons, nous nous proposons d'étudier plus concrètement les différentes mesures caractérisant les CR et les maillons de manière diversifiée. Le choix des mesures prend une place importante puisqu'il permet de dégager les CR dans le texte et de comprendre le phénomène sous plusieurs angles.

4.1 Les paramètres et leur signification

Dans cette partie, nous énumérerons la liste des paramètres couramment employés²⁴ et expliciterons, le cas échéant, leur signification dans des recherches des CR.

Parmi les paramètres caractérisant les chaînes de référence, nous retrouvons :

- Le nombre de mots : Dans l'exemple (17), la phrase se constitue de 13 mots et au sein desquels 4 référents apparaissent :

(17) [_{RF1} Cet été], [_{RF3} [_{RF2} votre] mission], si [_{RF2} vous] [_{RF3} l']acceptez : [_{RF2} vous] passer de [_{RF4} la climatisation] ! — (*Le Monde*, 01/08/2022)²⁵.

Par le nombre de mots du texte, nous pouvons calculer « la quantité d'éléments référentiels par rapport à la quantité de texte » (Boudreau & Kittredge, 2005 : 43)²⁶, ce qui correspond à 30,7 % dans ce cas précis. Nous pouvons également calculer la quantité de CR par rapport à celle de texte, ce qu'on appelle respectivement « la densité référentielle » et « la densité en CR » (Nous en reparlerons dans les paragraphes *infra*).

- Le nombre d'expressions référentielles du texte : il s'agit du nombre total de toutes les expressions référentielles existant dans le texte. Comme nous venons de le mentionner

²⁴ De nombreuses recherches sélectionnent une liste de divers paramètres en fonction de leur nécessité. Dans cette partie, la liste est basée principalement sur Schnedecker (2019 : 7-8) et Oberlé *et al.* (2018 : 92-93).

²⁵ « Canicule : les gestes pour réduire la température chez soi sans climatisation »
https://www.lemonde.fr/des-solutions-pour-la-planete/article/2022/08/01/canicule-les-gestes-pour-reduire-la-temperature-chez-soi-sans-climatisation_5480480_5444641.html

²⁶ Nous reprenons la citation à Schnedecker (2021 : 88).

dans le paragraphe précédent, ce paramètre sert à calculer la densité référentielle du texte. Dans la phrase (17) ci-dessus, 7 expressions référentielles renvoyant aux 4 référents sont observées.

- La densité référentielle : comme nous l'avons déjà mentionné dans les deux paragraphes précédents, la densité référentielle est calculée en divisant le nombre d'expressions référentielles par le nombre de mots du texte. Si nous l'appliquons à l'exemple (17), la densité référentielle est évaluée à 53,8 %.
- Le nombre de chaînes dans le texte : d'après Obry *et al.* (2017 : 96), le nombre de CR donne une indication sur le caractère mono- ou pluri-référentiel du texte. Lorsqu'un texte est mono-référentiel, il a une CR longue qui couvre tout le texte. C'est le cas de l'exemple (18) qui est un extrait du portrait journalistique, centré sur le personnage principal :

(18) **Charles Aznavour**, de **son** vrai nom Shahnourh Varenagh Aznavourian, est mort lundi 1^{er} octobre, à l'âge de 94 ans, dans **son** appartement situé dans les Alpilles (Bouches-du-Rhône), ont annoncé **ses** attachées de presse. En soixante-douze ans de carrière, **Charles Aznavour** a écrit plus de 1 400 chansons et en **a enregistré** plus de 1 200, chantées dans huit langues différentes. **Il** a également joué dans plus de quatre-vingts films et téléfilms. — (*Le Monde*, 01/10/2018)²⁷.

Pourtant, la plupart des textes sont pluri-référentiels, comportant plusieurs CR. Nous reprenons l'exemple (5) *supra* qui en comporte deux, celles d'« Ursus » et de « Homo », et donc rend le texte pluri-référentiel :

(5) **Ursus** et *Homo* étaient liés d'une amitié étroite. **Ursus** était un homme, *Homo* était un loup. Leurs humeurs s'étaient convenues. **C'**était l'homme **qui** avait baptisé *le loup*. Probablement **il s'**était aussi choisi **lui-même son** nom ; ayant trouvé Ursus bon pour **lui**, **il** avait trouvé *Homo* bon pour *la bête*. — (extrait du roman « L'Homme qui rit » par Victor Hugo, tome VIII., p.7).

- La longueur des chaînes : la longueur d'une CR s'obtient en comptant le nombre de maillons la composant. Une chaîne longue indique qu'il s'agit d'un référent important,

²⁷ « Charles Aznavour, un fils de l'exil parvenu en haut de l'affiche »

https://www.lemonde.fr/culture/article/2018/10/01/charles-aznavour-chanteur-et-comedien-est-mort_5362868_3246.html

par exemple, dans le roman, les CR des personnages principaux sont les plus longues. Dans l'exemple (19), qui est un extrait d'un roman, la chaîne la plus longue est celle sur le personnage « lord Clancharlie », avec 12 maillons. En revanche, la chaîne de l'« acceptation des pairs » n'ayant que 3 maillons est une chaîne courte et 'éphémère'.

(19) Il y avait dans ces temps-là un vieux souvenir. Ce souvenir était **lord Linnœus Clancharlie**.

Le baron Linnœus Clancharlie, contemporain de Cromwell, était un des pairs d'Angleterre, peu nombreux, hâtons-nous de le dire, qui avaient accepté la république. *Cette acceptation* pouvait avoir sa raison d'être, et s'explique à la rigueur, puisque la république avait momentanément triomphé. Il était tout simple que **lord Clancharlie** demeurât du parti de la république, tant que la république avait eu le dessus. Mais après la clôture de la révolution et la chute du gouvernement parlementaire, **lord Clancharlie** avait persisté. Il était aisé **au noble patricien** de rentrer dans la chambre haute reconstituée, les repentirs étant toujours bien reçus des restaurations, et Charles II étant bon prince à ceux qui revenaient à lui ; mais **lord Clancharlie** n'avait pas compris ce qu'on doit aux événements. Pendant que la nation couvrait d'acclamations le roi reprenant possession de l'Angleterre, pendant que l'unanimité prononçait son verdict, pendant que s'accomplissait la salutation du peuple à la monarchie, pendant que la dynastie se relevait au milieu d'une palinodie glorieuse et triomphale, à l'instant où le passé devenait l'avenir et où l'avenir devenait le passé, **ce lord** était resté réfractaire. **Il** avait détourné la tête de toute cette allégresse ; **il s'**était volontairement exilé ; pouvant être pair, **il** avait mieux aimé être proscrit ; et les années s'étaient écoulées ainsi ; **il** avait vieilli dans cette fidélité à la république morte. — (extrait du roman « L'Homme qui rit » par Victor Hugo, tome VIII., p.157)²⁸.

- La portée des chaînes : la notion de *portée* dans le domaine des CR est différente de celle de Charolles (1988) même si, à la base, elles proviennent de la même idée. Ici, nous parlons du « pourcentage de texte couvert par chaque chaîne », conformément à Schnedecker (2019 : 8), pour savoir si une chaîne est locale ou si elle s'étend à tout le texte. Reprenons les deux chaînes de l'exemple (19) *supra*. La chaîne de « lord Clancharlie » débute à partir de la deuxième phrase et continue tout au long du texte

²⁸ Wikisource.

jusqu'à la dernière phrase. Il s'agit d'une chaîne à longue portée, alors que, dans le cas contraire, à savoir une chaîne locale, la CR est considérée comme ayant une courte portée. C'est le cas de la chaîne de l'« acceptation des pairs », dont les trois maillons se placent dans une même phrase, et n'apparaît plus dans le contexte ultérieur.

- La distance inter-maillonnaire : Schnedecker (1997 : 11) souligne le rôle de la distance inter-maillonnaire en disant que plus les maillons sont proches, « plus la chaîne semble avoir de substance mais aussi de solidité ». Reprenons une partie de l'exemple choisi par Schnedecker (1997 : 10) où les deux premiers maillons « Michel Strogoff » ont 12 lignes de distance alors qu'il n'y a que 5 mots entre le deuxième maillon et le troisième²⁹.

(20) En homme qui ne craint ni le froid ni la neige, **Michel Strogoff** eût préféré voyager par la rude saison d'hiver, qui permet d'organiser le traînage sur toute l'étendue du parcours. Alors les difficultés inhérentes aux divers genres de locomotion sont en partie diminuées sur ces immenses steppes nivelées par la neige. Plus un cours d'eau à franchir. Partout la nappe glacée sur laquelle le traîneau glisse facilement et rapidement. Peut-être certains phénomènes naturels sont-ils à redouter, à cette époque, tels que permanence et intensité des brouillards, froids excessifs, chasse-neige longs et redoutables, dont les tourbillons enveloppent quelquefois et font périr des caravanes entières. Il arrive bien aussi que des loups, poussés par la faim, couvrent la plaine par milliers. Mais mieux eût valu courir ces risques, car, avec ce dur hiver, les envahisseurs tartares se fussent de préférence cantonnés dans les villes, leurs maraudeurs n'auraient pas couru la steppe, tout mouvement de troupes eût été impraticable, et **Michel Strogoff** eût plus facilement passé. Mais **il** n'avait à choisir ni **son** temps ni **son** heure. Quelles que fussent les circonstances, **il** devait les accepter et partir. — (extrait du roman « Michel Strogoff » par J. Verne, cité par Schnedecker, 1997, 10)

Ainsi, la distance entre les maillons peut fournir l'information sur le degré de cohésion

²⁹ Schnedecker (1997) synthétise trois façons de mesurer l'intervalle qui sépare deux maillons : typographiques, morphologiques et configurationnelles. Les approches typographiques comptent « le nombre de caractères ou de lignes entre deux mentions » et les approches morphologiques concernent « le nombre de syllabes (B.Lust, 1981) ou de mots (A. M. Gernsbacher, 1989 ; A. M. Gernsbacher & D. J. Hargreaves, 1988 ; Umstead & Leonard, 1983, cites par G. De Weck, 1991 : 72) ». En dernier lieu, la façon configurationnelle prend en compte « le nombre de phrases (M. Ariel, 1990 ; H. H. Clark & C. J. Sengul, 1979; T. Givon, 1983, 1989) ou de paragraphes (M. Ariel) » (*ibid* :35).

de la chaîne ou sur l'importance du référent dans le contexte en cours. Dans l'exemple (20), la faible distance inter-maillonnaire présentée à partir de l'occurrence du deuxième « Michel Strogoff » montre que cette chaîne est plus cohésive dans cette partie par rapport au début du texte.

- La longueur des maillons : la longueur des maillons peut varier d'un mot jusqu'à quelques lignes. Dans l'exemple (21) *infra*, le personnage « Joséphine Baker » est repris par les maillons de diverses longueurs. Par exemple, nous observons le maillon constitué d'un seul mot comme « elle » ou celui d'un ensemble de plusieurs mots comme « la fille d'une blanchisseuse et d'un père inconnu ». Il se pourrait également que le maillon ait une longueur de plusieurs lignes dans le cas d'un SN expansé comme par exemple « le gros chien que nous avons vu hier dans le parc près de chez moi qui allait fouiller mon sac pour voler mon sandwich jambon-beurre... ».

(21) Peu de personnalités incarnent aussi brillamment, aussi simplement, que **Joséphine Baker**, les valeurs que la France républicaine entend porter. En quittant, en 1925, à l'âge de 19 ans, **sa** ville natale de Saint-Louis (Missouri) pour Paris, **la fille d'une blanchisseuse et d'un père inconnu** fuit la ségrégation raciale et la pauvreté pour le pays de l'égalité. Toute **son** existence, aussi bien la façon dont **elle se** joue des clichés coloniaux sur la scène du Théâtre des Champs-Élysées, celle des Folies-Bergère ou celle du Casino de Paris, que **sa** vie sentimentale ou **son** engagement dans la Résistance, porte la marque de la liberté. Sans compter l'idéal de fraternité universelle de **cette infatigable militante contre le racisme** et de **cette mère d'une « tribu arc-en-ciel » de douze enfants**, adoptés aux quatre coins du monde. — (*Le Monde*, 26/08/2021)³⁰

- La composition des CR : les maillons peuvent être réalisés par différentes catégories grammaticales. Dans la chaîne de « Joséphine Baker » de l'exemple (21) *supra*, figurent le nom propre (*Joséphine Baker*), le déterminant possessif (*sa, son*), les SN définis (*la fille d'une blanchisseuse et d'un père inconnu*) et démonstratifs (*cette infatigable militante contre le racisme, cette mère d'une « tribu arc-en-ciel » de douze enfants*) et le pronom personnel (*elle*).
- Le mode de cohabitation des chaînes les unes avec les autres : comme nous l'avons

³⁰ « La leçon de France de Joséphine Baker »

https://www.lemonde.fr/idees/article/2021/08/26/la-lecon-de-france-de-josephine-baker_6092397_3232.html

mentionné dans la section 1.2 de ce chapitre, plusieurs chaînes cohabitent dans un texte de manières différentes : « succession, entrecroisement, dérivation, partition, fusion, déroulement parallèle. » (Schnecker, 1997 : 13-15).

- Les unités typo-dispositionnelles qu'elles couvrent : comme nous en avons parlé dans la section 1.1 au début de ce chapitre en reprenant Schnecker (2021 :173-184), les divers découpages typo-dispositionnels peuvent affecter les chaînes de sorte qu'ils « coïncident de manière assez systématique avec certaines catégories d'expressions référentielles » (*ibid* : 174). Les chapitres, les paragraphes, les titrilles, les discours rapportés ou même le parenthésage et les tirets peuvent servir à borner les CR.
- Le coefficient de stabilité : comme il a été mentionné dans la section 3.2 de ce chapitre, cette mesure proposée par Perret (2000 : 17) concerne la diversité lexicale des expressions référentielles employées désignant le même référent. Reprenons l'exemple (21) *supra*, « Joséphine Baker » est désignée par les termes suivants : *Joséphine Baker, la fille d'une blanchisseuse et d'un père inconnu, cette infatigable militante contre le racisme, cette mère d'une « tribu arc-en-ciel » de douze enfants.* Dans ce cas-là, les quatre anaphores nominales désignant le personnage sont toutes différentes et la chaîne n'est pas stable.

En règle générale, parmi ces divers paramètres caractérisant les CR, seuls ceux qui conviennent à l'objectif et aux besoins d'une recherche sont choisis pour être analysés. La partie suivante synthétisera les fonctions de chaque mesure, en vue de construire plus aisément une stratégie d'analyse.

4.2 La fonction des paramètres dans l'analyse

Oberlé *et al.* (2018 : 92-93) sous-catégorisent les indicateurs en trois selon leur fonction : 1) ceux qui révèlent les caractéristiques référentielles du texte, 2) ceux qui décrivent les caractéristiques des CR et 3) ceux qui montrent la relation entre les chaînes et la structure du texte.

De la même façon, Schnecker (2021 : 109) classe quatre types de mesures caractérisant les CR. Ce classement nous permet de voir, d'emblée, ce qu'un paramètre qui exploite certaines données brutes révèle. En d'autres termes, elle montre la corrélation entre les paramètres, les données brutes et, de plus, ce que nous pouvons savoir à partir de ces données. Nous reprenons son tableau récapitulatif pour construire la stratégie d'analyse :

	Paramètres	Ce qui est compté / calculé / observé	Ce qui est révélé
Densité référentielle	Ratio ER / mots	Nbre d'ER	Opposition diamésique (oral vs. écrit), genres de discours
Densité en CR	Nbre de CR	Nbre de référent faisant l'objets d'une CR	Nbre et importance respective du/des thème(s) discursif(s)
Dimensions des CR	Longueur	Nbre de mentions d'une chaîne	Importance du référent
	Distance inter-maillonnaire	Nbre de mots, tokens, etc. entre deux mentions	Cohésion de la chaîne et degré d'importance du référent
	Portée / étendue	Nombre de linges, §, etc. que couvre une chaîne	Degré d'importance du référent
Unités d'occurrence des CR	Occurrence dans :		
Propositionnelle	Prop.matrice vs. subordonnée	Nbre d'occurrences et modalités de désignation d'un référent/unité	Saillance du référent
Textuelle	Paragraphe, section/chapitre, etc.		Importance du référent et hiérarchisation des référents
Sémantico-discursive	Domaine (spatiaux, temporels, fictionnels, etc)		Phases/facettes (voire évolution) des référents
	Point de vue		
Caractéristiques des maillons	Catégorie grammaticale des maillons	Nbre d'occurrence de chaque cat.gram	Cohésion de la chaîne Rapport/découpages textuels et discursifs
	Catégorie lexicale des maillons	Nbre d'occurrence de chaque ss-cat. nominale/CR	Importance relative des référents Genres textuels
	Longueur des maillons	Nbre de mots, tokens, etc./maillon	Cohésion de la chaîne Genres discursifs d'occurrence
	Complexité syntaxique des maillons	Nbre et type de modificateurs /maillons	
	Coefficient de stabilité	Variation de la tête lexicale des mentions	Genres discursifs, date, et système linguistique de composition
	Fonction syntaxique des maillons	Nbre d'occurrences par fonction	Saillance référentielle
	Rôle sémantique des maillons	Nbre d'occurrences par rôle thématique	
Modalités de cohabitation	Successions /	Répartition des mentions des	Développement des thèmes de

des CR	entrecroisement	CR respectives les unes par rapport aux autres	discours dans le texte
	Fusion / partition		
	Parallélisme		

Tableau 4 Synthèse des paramètres caractérisant les CR (Schneidecker, 2021 : 109)

Comme nous l'avons déjà mentionné, ces mesures ne sont pas forcément toutes employées dans le travail d'analyse des CR. Pourtant, il est important de bien comprendre chaque paramètre, en raison de son apport pour l'appréhension des particularités des chaînes et de la structuration du texte.

Bilan

En guise de synthèse, nous allons préciser notre position sur la définition de *CR*, notamment sur quelques points obscurs à propos de la délimitation des *CR*.

La question sur les bornes de la chaîne est un sujet important (voir la section 1.1 de ce chapitre). Contrairement aux anaphores et coréférences qui se présentent souvent en une phrase ou dans des phrases juxtaposées, les *CR* s'étendent sur tout un texte. De ce fait, la notion de *CR* concerne les processus cognitifs du lecteur, et surtout, dans des textes longs comme un roman de plusieurs centaines de pages, le problème du coût de traitement se pose.

Comme nous l'avons déjà mentionné dans la section 1.1 *supra* en reprenant l'idée présentée dans Schneidecker (2021), les *CR* peuvent être délimitées de manières diverses : découpages typographiques, découpages sémantiques et découpages marqués par les expressions référentielles. Malgré la validité de cette idée proposée par Schneidecker (1997, 2021), nous allons considérer qu'une chaîne couvre tout le texte sans coupure. Ce choix sera détaillé lorsque nous parlerons de la méthodologie y compris de la présentation du corpus.

En synthétisant les caractéristiques des maillons, nous avons aussi parlé de la catégorie grammaticale des expressions référentielles. Comme nous l'avons souligné dans la section 2.1 de ce chapitre, le consensus sur la délimitation des expressions référentielles qui sont susceptibles de servir de maillon n'est pas encore atteint. Notre position sur le sujet sera traitée dans le chapitre 4.

Ensuite, les trois facteurs qui influencent la variété de matériau lexical sont explicités : le

genre textuel, la date de la composition et la langue de rédaction du texte (voir la section 3). Parmi ces facteurs mentionnés, notre étude se concentre notamment sur la langue de la rédaction et vise à comparer les CR dans le texte français et dans le texte coréen.

Pour finir, un ensemble des mesures pour analyser les CR est présenté à la fin de ce chapitre. Ces divers paramètres, qui servent à quantifier les CR dans le texte, montrent la complexité du phénomène (voir le tableau 4 ci-dessus). Certains d'entre eux seront employés lors de l'analyse de notre corpus dans le chapitre 7.

Chapitre 3.

Les chaînes de référence d'un point de vue contrastif

Introduction

Appliquer le point de vue comparatif, quel que soit le domaine (le genre textuel, la langue, l'époque de la rédaction, l'écrit ou l'oral), est une façon efficace de mener une recherche linguistique, la comparaison des différents facteurs permettant d'observer plus facilement les points universels et individuels du phénomène en question.

Notre idée de départ est de réaliser une étude inter-langues. Cette approche des CR qui tient compte d'une des trois conditions concernant la variété du matériau lexical (le genre textuel, la date de rédaction et la langue de rédaction, voir § 3 du chapitre 2) pourrait donner un aperçu synthétique sur le phénomène linguistique et, en même temps, un point de vue contrastif sur les particularités de chaque langue.

Ce troisième chapitre est donc consacré, dans un premier temps, à l'état de l'art des approches inter-langues dans le champ linguistique, notamment celui des CR, du français et du coréen (§1). Dans un second temps, nous présenterons quelques particularités du coréen à prendre en compte (§2) avant d'entamer l'analyse des expressions référentielles et des CR de chaque langue. Les descriptions comparatives fournies dans les chapitres ultérieurs (4 et 6 entre autres) concerneront toute ou une partie de ce que nous expliciterons dans ce présent chapitre.

1 Approche inter-langues : comment refléter les différences ?

La linguistique française, qui a commencé à s'intéresser aux CR relativement tôt par rapport à la linguistique coréenne, comprend un certain nombre de recherches basées sur le point de vue inter-langues. En revanche, la linguistique contemporaine de la langue coréenne, qui n'a pris corps qu'à partir du début du 20^{ème} siècle en raison de l'occupation japonaise et de sa politique visant à faire disparaître le coréen, n'est pas encore suffisamment avancée dans la recherche sur les CR.

Partant de ce constat, dans ce sous-thème, nous ferons un bref tour d'horizon des études de nos prédécesseurs : nous aborderons directement les travaux inter-langues de la linguistique française traitant le problème du CR, et, pour la linguistique coréenne, nous présenterons les recherches inter-langues sur les notions connexes de CR. Cette démarche permettra de se faire une idée globale de l'approche inter-langues qui sera appliquée dans le cadre de cette thèse.

1.1 Approche inter-langues dans la linguistique française

Avec l'intérêt croissant des linguistes sur les différentes branches de la référence telles que la coréférence, l'anaphore et les CR, l'approche inter-langues commence à être exploitée (les articles dans la revue *Travaux de linguistique*¹, entre autres). Nous allons évoquer quelques études sur les CR choisissant l'analyse d'un point de vue comparatif entre deux langues dont l'une des langues comparées est le français, à part les recherches déjà mentionnées plus haut (voir la section 3.3 du chapitre 2), à savoir l'étude de Lundquist (2005) et celle de Baumer (2015).

Baumer, Dias & Schnedecker (2020) travaillent sur un corpus trilingue (anglais, allemand et français) de 45 textes constitués d'*incipit*² romanesques réalistes du 19^{ème} siècle. Les auteurs essaient de montrer les différences inter-langues *via* la composition des CR de chaque langue, à savoir la proportion des expressions référentielles employées, et présentent le schéma instancié pour l'introduction des personnages principaux. De telles disparités ne

¹ N° 82, paru en 2021.

² Environ 500 premiers mots des romans.

peuvent pourtant pas être explicitées par le type de langue. Les auteurs admettent que, dans ce corpus, la question du genre textuel semble un facteur important.

Baumer *et al.* (2021) examinent aussi un corpus trilingue, anglais, allemand et français, mais cette fois-ci, il est constitué des nouvelles d'actualité extraites de la presse. Les auteurs soulignent que les CR dans les textes des trois langues sont communément influencées par la structure générale de l'information. Cependant, une différence importante est remarquée entre les textes anglais et les textes écrits dans les deux autres langues : les informations qui n'ont pas de lien immédiatement visible sont souvent juxtaposées en anglais tandis que le français et l'allemand montrent plus de continuité à cet égard. Cette différence mérite d'être soulignée parce qu'elle ne peut pas être explicitée par la distinction entre les langues romanes (français) et les langues germaniques (anglais et allemand).

L'étude inter-langues français et chinois de Guo (2021) sur des textes encyclopédiques sur les animaux analyse la composition des CR et la cohabitation des CR dans le texte. En vertu de la particularité du genre textuel, les deux langues montrent quelques similarités telles que le composant principal des CR (les groupes nominaux sans modifieur) et le mode de cohabitation des CR³. En revanche, au niveau de l'emploi des pronoms, une disparité importante est observée : « les pronoms et les déterminants possessifs sont plus souvent utilisés que les pronoms zéro en français, tandis que le phénomène inverse est observé en chinois » (*ibid*, 149).

Aborder le phénomène linguistique d'un point de vue inter-langues nous oblige à prendre en considération les caractéristiques des langues dans leur globalité. Toutefois, comme nous venons de le voir, selon le genre traité comme corpus, l'objectif de recherche ou encore l'intérêt de l'auteur, l'approche inter-langues peut choisir de couvrir soit l'ensemble des différences typologiques des langues concernées, soit uniquement un ou plusieurs aspects spécifiques parmi ces différences.

Il est évident qu'une réflexion globale et profonde sur la langue elle-même et sur les manières dont chaque langue aborde le problème de la référence doit constituer un préalable, afin de bien comprendre le phénomène des CR qui se présente différemment selon la langue.

³ D'après l'auteur (Guo, 2021), dans les deux langues, la majorité des paragraphes sont mono-référentiels. Dans les cas de paragraphes pluri-référentiels, l'entrecroisement et la dérivation sont les deux modes de cohabitation les plus souvent observés.

1.2 Approche inter-langues dans la linguistique coréenne

La notion de *CR* étant encore inconnue dans la linguistique coréenne, il existe très peu de recherches sur cette notion, et peu d'études ont été menées du point de vue contrastif inter-langues. Par rapport au problème de la référence, il n'y a guère plus : nous ne trouvons que quelques recherches inter-langues sur les expressions référentielles, surtout sur les démonstratifs ou les expressions ostensives.

Certains auteurs d'origine chinoise ou les enseignants coréens en Chine mènent des études sur les démonstratifs en chinois et en coréen. Hu J. (2011) analyse les fonctions déictiques des démonstratifs dans les deux langues : deixis contextuelle, anaphorique et conceptuelle⁴. Zhu W. (2018) analyse également les emplois des démonstratifs en coréen pour proposer quelques stratégies didactiques aux étudiants chinois apprenant le coréen.

Les études comparatives sur le japonais et coréen suivent la même tendance (Bae D. H. : 1998, Seo N. R. : 2014). Les auteurs travaillent sur les différents usages des démonstratifs de chaque langue et visent à les appliquer à la didactique des langues.

La plupart de ces recherches inter-langues ne dépassent pas les bornes de la comparaison ponctuelle des expressions démonstratives. Cette tendance qui pourrait contribuer aux domaines de la traduction et de la didactique des langues n'est cependant pas suffisante pour aborder une étude sur un phénomène spécifique de référence tel que les *CR*. Nous allons donc examiner quelques études comparatives menées sur le coréen, ainsi qu'une autre portant sur des langues occidentales.

D'abord, l'approche inter-langues anglais-coréen de Yang Y. J. (2018) tente d'examiner les caractéristiques des déterminants démonstratifs des deux langues, « this / that / it » en anglais et « i / jeo / geu » en coréen, et de comparer leurs emplois.

Une autre étude sur l'anglais et le coréen de Song J. Y. (2011) analyse le système des déterminants démonstratifs. L'auteur se concentre sur le fait qu'en anglais le classement des démonstratifs est binaire (proximal et distal) alors qu'en coréen, il est ternaire : en plus des démonstratifs proximal et distal, le coréen distingue le démonstratif dont le pivot de la référence est le locuteur de celui dont le pivot est l'interlocuteur. L'objectif de cette étude est

⁴ Il s'agit d'une notion qui désigne l'acte de renvoyer à un objet se situant dans la représentation mentale du locuteur et de son interlocuteur, en l'occurrence, le référent n'est physiquement manifesté ni dans le texte ni dans le contexte.

de trouver un compromis entre les deux systèmes afin d'examiner la manière dont on traduit des démonstratifs anglais en coréen.

Ahn J. K. & Eun H. Y. (2006) analysent également quelques techniques de traduction des démonstratifs anglais en coréen. Selon eux, en raison de la haute dépendance au contexte que manifeste la langue coréenne, la comparaison inter-langues des expressions démonstratives nécessite de tenir compte de l'aspect pragmatique.

Ainsi, de nombreuses recherches inter-langues basées sur l'expression référentielle en coréen se focalisent sur le démonstratif. Cette tendance explique le manque d'intérêt de la linguistique coréenne pour la notion de *référence* et diverses autres expressions référentielles. Le démonstratif, en tant qu'expression référentielle la plus représentative en raison de son caractère ostensif, est facilement considéré comme objet de comparaison.

Il existe aussi quelques recherches sur le français et le coréen. Shim B. S. (1995) analyse la forme du SN démonstratif « ce + N » en français et en coréen. Malgré le fait que le français et le coréen appartiennent à deux familles linguistiques très différentes l'une de l'autre, du point de vue syntaxique, il souligne que les deux SN se ressemblent par leur fonction, leur sens lexical et leur comportement anaphorique.

Hong Y. T. (2009) compare les catégories fonctionnelles des SN du coréen avec celles du français du point de vue syntaxique, plus précisément, du point de vue de la grammaire générative. Les deux langues étant complètement différentes, l'auteur analyse les éléments susceptibles de constituer un SN et leur ordre d'apparition dans le SN. Il insiste sur le fait que, contrairement au français, le coréen n'a pas de phénomène d'accords (ni nombre ni genre) et que la marque de pluralité « -deul »⁵, ainsi que la particule de cas qui s'ajoute à la fin du mot prennent une place importante dans la syntaxe des SN en coréen.

Park D. Y. (2001) parle du problème de la détermination du nom en français et en coréen. En tenant compte du fait que le système de l'article est absent en coréen, il insiste sur le fait que le système de la particule en coréen peut, à certains égards, faire office de déterminant. Dans cette perspective, l'auteur analyse quelques particules en coréen et leurs emplois d'un point de vue psychomécanique. Selon lui, même s'il est difficile de dire que la particule en coréen et l'article en français se correspondent, l'article indéfini « un » et la particule « -i / -ga » ont un point commun dans le sens où ils participent à l'actualisation du mot. Par contraste,

⁵ La marque de pluralité « -deul » postpose au nom afin d'exprimer la pluralité du nom en question.

l'article défini « le » et la particule « -eun / -neun »⁶ montrent le procès de virtualisation du mot.

Malgré ces diverses approches inter-langues des expressions référentielles, nous admettons que la linguistique coréenne reste encore assez indifférente au problème de CR⁷. Notre étude sera donc l'une des premières qui fournira à la fois une vue globale sur le phénomène et une analyse approfondie sur les différences entre le français et le coréen. Les recherches antérieures que nous venons de mentionner, même si elles ne travaillent pas précisément sur les CR, sous-tendront notre analyse des CR en coréen.

2 Les particularités du coréen à considérer dans le domaine de la référence

Dans les pages suivantes, nous évoquerons quelques caractéristiques importantes du coréen à prendre en considération tout au long de cette étude. Déclinée en trois sous-sections, cette partie se concentre d'abord sur la description générale de la langue et sur ses deux particularités les plus notables qui sont susceptibles d'affecter les CR. Cette étape nous permettra d'introduire les chapitres 4 et 6, consacrés aux expressions référentielles dans les deux langues et aux stratégies d'annotation.

2.1 Description générale de la langue coréenne : une langue agglutinante

Le coréen⁸ est généralement considéré comme un isolat⁹ qui ne montre aucune relation

⁶ Pour la notion de *particule*, son classement et sa fonction dans la langue coréenne, voir la section 2.2.1 dessous.

⁷ Jusqu'à présent, la notion n'est pas suffisamment traitée en Corée. Un seul article sur la CR intitulé « The Triangle of Referential Chains : Anaphor, Association and Intuitive equi-reference » est publié en 2017 par Kim K. S. En outre, nous nous permettons d'évoquer notre participation aux colloques SCELLF en 2019 (communication intitulée « Etude contrastive des chaînes de référence en français et en coréen ») et KAFEL en 2020 (« Comment annoter les chaînes de référence ? »).

⁸ Langue parlée en grande partie en péninsule coréenne y compris la Corée du Sud et la Corée du Nord, dans les quartiers frontaliers de la Chine (district Yanbian) et dans les communautés émigrées des coréens dispersées dans le reste du monde.

⁹ Velupillai (2012: 22) définit la notion d'« isolate » comme suit : « Some languages cannot be demonstrated to have any genealogical relationship with any other language. These kinds of languages, called language isolates,

génétiq ue avec d'autres langues. Par ailleurs, selon les travaux de typologues appuyés sur la morphologie, le coréen est une langue agglutinante¹⁰ enclitique : les éléments linguistiques comme les particules, l'affixe et la terminaison qui portent des informations sur le rôle ou sur la valeur grammaticale d'un mot, sont attachés à la racine ou au radical du mot qui représente le sens. L'exemple suivant montre le fonctionnement de ce système :

- (1) 강아지가 사과를 먹었다.¹¹
gangaji+ga sagwa+leul meog+eot+da.¹²
chiot+cas-sujet pomme+cas-objet manger+passé+terminaison
 Traduction en français : Le chiot a mangé la pomme.

Dans cette phrase, les premiers deux mots surlignés « 강아지 (gangaji) / chiot » et « 사과 (sagwa) / pomme » sont des substantifs qui gardent toujours leur forme invariable. Par contre, « -가 (-ga) » et « -를 (-leul) » sont des particules présentant les rôles syntaxiques de « 강아지 (gangaji) » qui fonctionne comme sujet, et « 사과 (sagwa) » qui est l'objet de la phrase. Ils ne s'emploient pas tous seuls et se postposent toujours à un autre mot. Ensuite, le troisième élément souligné « 먹 (meog) / manger » est la racine du verbe « manger ». L'élément suivant « -었 (-eot) » est l'affixe qui porte l'information sur le temps et l'aspect qui est ici le passé, et « -다 (-da) » qui se trouve à la fin est la terminaison du verbe d'une phrase déclarative.

Le coréen étant une langue agglutinante qui se définit par le fait qu'un élément morphologique s'ajoute à un élément lexical, la linguistique coréenne accorde une grande importance à la morphologie, tandis que la linguistique textuelle reste un domaine moins approfondi. Dans cette partie, nous examinerons quelques caractéristiques de la langue

can be thought of as belonging to a family which is made up of only one genus, which in turn consists of only one language. ».

¹⁰ À part le coréen, plusieurs langues, le japonais, le mongol, le turc, etc., sont classées comme langues agglutinantes.

¹¹ Pour la translittération, nous utilisons la manière la plus souvent utilisée dans les études comparatives français-coréen : Kim S. Y. (2010), Son H. (2005).

¹² Nous utilisons le système de la transcription du coréen en alphabet latin défini par l'Académie nationale de la langue coréenne (<https://kornorms.korean.go.kr/>). Le tableau qui montre la correspondance entre les phonèmes coréens et l'alphabet latin figure dans la page de *Règle de translittération* (partie préliminaire au début de la thèse). Pour les mots d'origine étrangère, nous respectons l'orthographe originale.

coréenne (comprenant le trait morphologique qui vient d'être mentionné) afin de dégager certains points du coréen qui pourront concerner le phénomène des CR et leurs composants.

Yoo H. K. *et al.* (2015 : 34-36) définissent les caractéristiques du coréen en comparaison avec les langues altaïques¹³. Nous faisons le tri dans leur liste en extrayant certaines caractéristiques qui sont aussi comparables avec le français et en excluant celles qui ne sont pas considérées comme pertinentes dans le domaine de la référence et des expressions référentielles. Cela étant, certains traits phonologiques ou syntaxiques ne sont pas compris dans la liste suivante :

- l'existence des éléments morphologiques qui sont postposés aux noms (particule, affixe et terminaison),
- les qualificatifs qui se placent toujours devant les noms,
- l'absence de pronoms relatifs,
- l'absence de marqueurs de la personne.

Ces traits seront mentionnés *via* nos réflexions dans les pages suivantes, et plus particulièrement dans le chapitre 4, lorsque nous parlerons des expressions référentielles de chaque langue. Pour l'instant, sans entrer dans les détails de ces traits, nous allons évoquer les deux grandes tendances caractéristiques du coréen, qui le distinguent de la langue française, à savoir l'absence de l'article et la haute dépendance de contexte.

2.2 L'absence de l'article et l'existence d'un système indirect et allusif

Le coréen ne possède pas de système d'article. L'article, en règle générale, montre le type de relation formelle et référentielle qu'un nom établit avec d'autres entités dans la phrase. Autrement dit, il permet de voir les caractéristiques du référent réalisées dans le contexte, telles que le genre, le nombre et les nuances sémantiques, auxquelles renvoie le nom en question. Cependant, malgré l'absence de l'article qui fonctionne comme indice quantificatif ou référentiel du nom, les coréénophones n'ont aucune difficulté à discerner le caractère défini

¹³ Les langues altaïques sont un groupe de langues parlées en Eurasie : de l'est de l'Europe jusqu'à l'Asie de l'est. Les langues turques (le turc, l'azéri, le turkmène, *etc.*), les langues mongoles (le mongol, l'oïrate, le kalmouk, *etc.*) et les langues toungouses (les langues parlées en Sibérie comme l'evenki, l'évène, le néguidale, *etc.*) appartiennent à cette famille de langue. Dans certaines conceptions, les langues coréaniques et les langues japoniques sont également considérées comme les langues altaïques mais cela reste contestable.

unités nominales en coréen. L'auteur insiste sur le fait que, malgré l'absence de système d'article déterminant immédiatement le nom, tous les noms et les SN possèdent une sorte de caractère défini ou indéfini, et ce, quel que soit leur statut référentiel : référentiel, non-référentiel, spécifique, générique. Pour ainsi dire, selon lui, la *détermination* est une notion fonctionnellement indispensable pour les noms et l'absence de l'article en coréen peut être indirectement comblée à la fois par le contexte et par d'autres constituants de la phrase¹⁷.

L'article est, en général, considéré comme l'élément linguistique le plus immédiat et le plus contraignant au niveau de l'attribution de définitude. Faute de système d'article, le coréen, en s'appuyant sur le contexte d'énonciation pour saisir la définitude de référent, choisit également d'autres éléments de la phrase qui peuvent se consacrer de façon indirecte à la réalisation de la détermination : la particule et le prédicat.

2.2.1 Rôle et fonctionnalité discursive de la particule en coréen

La particule en coréen est un élément linguistique complexe. Bien qu'elle soit dotée d'un trait de morphème et qu'elle ne soit jamais employée seule, la particule est considérée comme un mot dans la grammaire coréenne. Elle s'ajoute toujours à un autre élément indépendant, à savoir les substantifs, pour montrer le rôle de ces derniers et la relation qu'ils entretiennent avec les autres unités linguistiques de leur phrase d'accueil (Nam K. S. *et al.*, 2019 : 70). D'après cette définition canonique, il y a trois types de particules réparties selon leurs fonctions : la particule de cas, la particule auxiliaire et la particule conjonctive. Cependant nous ne parlerons que des particules de cas et d'auxiliaire en raison de leur participation à la détermination d'un mot, contrairement à la particule conjonctive.

La particule de cas postposée à un mot donne l'information sur la fonction de ce mot dans la phrase. Il existe plusieurs types de particules de cas pour les différents constituants de la phrase : sujet, objet, complément, prédicat, déterminant, adverbe et vocatif¹⁸. Reprenons à

¹⁷ Comme nous l'avons mentionné en reprenant l'idée de Hong C. S. (2001) dans le paragraphe précédent, la (in)détermination peut être immédiatement réalisée au niveau lexical, c'est-à-dire par diverses expressions épithètes. Dans cette partie, nous ne le mentionnons pas parce que notre objectif est de décrire les éléments grammaticaux non immédiats consacrés à la détermination d'un mot. Quant à l'aspect lexical, nous aurons l'occasion d'en reparler dans le chapitre 4. Par ailleurs, les expressions épithètes seront prises en compte dans le travail d'annotation en tant que déterminants.

¹⁸ Dans la linguistique coréenne, la classification des particules est le sujet le plus discuté dont la complexité empêche le consensus. Nous allons suivre la classification canonique de la forme des particules de cas comme le

nouveau l'exemple (1) :

- (1) gangaji+ga sagwa+leul meog+eot+da.
 chiot+cas-sujet pomme+cas-objet manger+passé+terminaison
 Traduction en français : Le chiot a mangé la pomme.

Comme nous l'avons brièvement expliqué dans la section 2.1 *supra*, dans cette phrase, le substantif « chiot » est accompagné de « -ga » alors que « pomme » est accompagné de « -leul ». Ainsi, la particule de cas sujet¹⁹ signale que le mot auquel elle est rattachée est sujet dans la phrase, celle d'objet que le mot auquel elle est rattachée est objet²⁰. Parmi les diverses

montre le tableau récapitulatif qui suit. Les différentes particules sont employées selon le rôle du constituant auquel elles s'ajoutent.

Type	Forme
Particule-sujet	-i / -ga *-kke-seo (l'emploi honorifique), -e-seo (accompagné avec un groupe ou une organisation), etc.
Particule-objet	-eul / -leul
Particule-complément attributif	-i / -ga (Cette particule a la même forme que la particule-sujet. Pourtant, elle ne s'emploie qu'avec certains verbes qui constituent une construction attributive tels que 'devenir', 'ne pas être', 'aimer', 'détester', etc.)
Particule-prédicat	-ida (être + N attribut)
Particule-possessive	-ui
Particule-adverbe	-e(à-lieu) / -ege(à-personne) / -wa(avec) / -boda(plutôt) / -lo(en), etc. (Les particules de ce type font fonctionner les substantifs avec lesquels elles s'emploient comme l'adverbe.)
Particule-vocatif	-a / -ya

Tableau récapitulatif de particule du cas

¹⁹ Les termes qui désignent les sous-catégories de la particule de cas sont nommés de plusieurs manières différentes en français. Par exemple, la particule de cas accompagnant le sujet de la phrase est appelée « la particule nominative » (Kwon & Zribi-Hertz, 2006), « la particule-sujet » (Son H., 2005), « la particule du sujet » (Yun H.S., 2004), etc. Cette disparité découle des différentes traductions de chaque auteur. Dans cette étude, nous allons suivre la dénomination de Son H. (2005), la particule-'le nom de constituant', (par exemple. « la particule-sujet »), afin de simplifier l'explication et de transmettre plus clairement le rôle de chaque particule.

²⁰ Normalement, grâce à l'existence de la particule de cas qui indique le rôle d'un mot joué dans la phrase, l'ordre des mots est moins important en coréen qu'en français. Cependant, la particule de cas peut être ellipsée si l'on connaît avec certitude la fonction et le sens du SN, par exemple, dans les cas où la phrase respecte bien l'ordre SOV :

- ex1) gangaji+ga sagwa+leul meog+eot+da.
 → gangaji sagwa meogeotda. (Le chiot a mangé la pomme.)

À l'oral, il se peut que la particule de cas ne se présente pas même dans le cas où l'ordre SOV n'est pas respecté :

- ex2) sagwa gangaji meogeotda.

formes de particule de cas (mentionnées dans le tableau de la note 22), nous allons traiter «-i / -ga »²¹ et « -eul / -leul »²² qui concernent immédiatement la constitution des SN fonctionnant comme un sujet et un objet.

À l'instar de la particule de cas, la particule auxiliaire s'agglutine à la fin du mot mais elle joue un rôle différent : cet élément grammatical correspond à ce qui serait une préposition dans le système français ou un élément conférant au mot un sens adverbial : « de », « jusque » ou bien « aussi », « même », « seulement », « chaque », *etc.*

Plus particulièrement, ce qui nous intéresse est la particule auxiliaire « -eun / -neun » qui est considérée comme ayant une double fonction : elle peut être utilisée non seulement pour donner une description « neutre » du nom auquel elle s'ajoute, mais aussi, selon les cas, pour signaler une construction contrastive, soit explicite, soit implicite :

(2) 강아지는 사과를 먹었다.

gangaji+neun sagwa+leul meog+eot+da.

chiot+cas-aux pomme+cas-objet manger+passé+terminaison.

Traduction en français :

(Description neutre) Le chiot a mangé la pomme.

(Construction contrastive) Le chiot a mangé la pomme alors que (p.e. le chat...)

Ici, nous avons substitué la particule sujet « -ga » de l'exemple (1) à la particule auxiliaire « -neun ». Si cette phrase se situe au tout début du discours sans autres indices contextuels, la particule « -neun » fonctionne comme un marqueur de topique, sinon, le SN « gangaji+neun » inclut la présupposition qu'il existe un objet de comparaison même s'il n'y a pas de proposition supplémentaire explicite. Dans ce dernier cas, en lisant cette phrase, le lecteur comprend tout de suite l'intention du locuteur, qui est de signaler le fait que le chiot a mangé la pomme, ainsi que la suite telle qu'un autre animal, quel qu'il soit, n'a pas mangé (ou n'a pas pu manger) la pomme.

pomme chiot a mangé. (Le chiot a mangé la pomme.)

²¹ Ce sont des allomorphes qui sont exclusifs l'un de l'autre et choisis selon la règle phonologique. Pour une raison de liaison, si le *chae-eon* (mot) qui précède la particule se termine par une voyelle, la particule-sujet doit être « -ga ». Et s'il se termine par une consonne, la particule est « -i ».

²² Ce sont aussi des allomorphes employés différemment selon la règle phonologique. Quand le *chae-eon* se termine par une voyelle, la particule-objet est « -leul ». En revanche, quand le *chae-eon* se termine par une consonne, ce mot est accompagné par « -eul ».

Cette définition de la particule étant canonique, les linguistes d'aujourd'hui essaient d'associer cette dernière avec le problème de la détermination. Plusieurs auteurs décrivent la différence de ces trois paires de particules qui peuvent participer à la détermination d'un nom en coopération avec d'autres éléments de la phrase. Commençons par les exemples simples :

(3) 강아지는 귀엽다

gangaji+neun gwiyeob+da.

chiot+particule mignon+terminaison.

Traduction en français : Le chiot (générique) est mignon.

(4) 강아지가 귀엽다.

gangaji+ga gwiyeob+da.

chiot+particule mignon+terminaison.

Traduction en français : Le chiot (spécifique) est mignon.

Ces deux exemples ont la même structure, excepté la particule agglutinée à « chiot ». Cependant, à condition que l'on ne présuppose pas d'autre indice contextuel, cette disparité n'est jamais triviale parce que le premier énoncé parle du chiot de manière générale²³, plus spécifiquement, il s'agit d'un référent dénotatif, alors que le chiot mentionné dans (4) désigne un référent concret²⁴.

²³ Comme nous l'avons mentionné en décrivant l'exemple (2), la particule « -neun » peut participer à la description neutre en tant que marqueur de topique.

²⁴ Bien évidemment, comme nous venons de l'expliquer avec (2), le chiot de (3) peut être interprété comme un référent concret si la phrase est dite dans un contexte d'énonciation tel que l'on est en train de regarder plusieurs animaux :

ex) (goyangi+neun an gwiyeoun+de) gangaji+neun gwiyeob+da.

chat+particule négation mignon+term-connexion-opposition chiot+particule mignon+terminaison.

Traduction en français : (Le chat n'est pas mignon mais) le chiot est mignon.

De même, le chiot de (4) est interprétable comme un référent générique dans le contexte d'énonciation où l'on parle de sa préférence ou sa passion.

ex) (yojeum) gangaji+ga gwiyeob+da.

Ces jours-ci chiot+particule mignon+terminaison.

Traduction en français : (Ces jours-ci, je sens que) le chiot est mignon. (Je ne le sentais pas auparavant.)

Ainsi, le contexte d'énonciation est une variable importante dans l'interprétation, mais, pour l'instant, nous nous concentrons uniquement sur le rôle théorique de la particule dans la détermination de nom.

La description de Cornish (2011 : 139) sur la proposition thétiq ue au niveau discursivo-pragmatique peut expliciter l'interprétation du « chiot » dans l'exemple (4). Selon lui, la raison d'être de la proposition thétiq ue « est de servir à présenter une entité, une proposition ou un état de choses en tant qu'élément d'information nouveau pour le discours ».

Ainsi, la plupart des auteurs soulignent que les particules « -i / -ga » et « -eul / -leul » sont thétiq ues, tandis que « -eun / -neun » sont catégoriq ues. Park H. S. (2014) met l'accent sur le rôle expressif de la particule : elle le présente en distinguant le rôle de la particule de cas (-i/-ga) de celui de la particule auxiliaire (-eun/-neun). La première peut être expliquée par le terme de *jugement thétiq ue*, à savoir un rôle de focalisateur, alors que la deuxième emprunte la notion de *jugement catégoriq ue*²⁵ qui a un rôle topical. Le tableau que nous proposons ci-après récapitule le propos de l'auteur :

Particule	Structure basique	Présupposition
-i / -ga	A (/ B / C...)+ga P	Le prédicat est figé dans la tête du locuteur : seul le sujet est sélectionné.
-eun / -neun	A+neun P (/ Q / R...)	Le sujet est figé dans la tête du locuteur : seul le prédicat est sélectionné.

Tableau 5 Structure sous-jacente de l'énoncé selon le choix de la particule

Choi J. J. (2016) décrit aussi cette différence entre les deux types de particule, en mettant en rapport avec les différents niveaux de la relation informationnelle du texte²⁶. Toutefois, cette distinction n'est pas suffisante pour rendre compte de la différence entre les cas suivants :

- (5) 말이 달린다.
mal+i dalri+n+da.
cheval+particule courir+présent+terminaison.

²⁵ Kuroda *et al.* (1973) présente la théorie du jugement originellement proposée par Brentano et élaborée par Anton Marty. Les notions de « jugement thétiq ue » et de « jugement catégoriq ue » sont respectivement appliquées aux particules japonaises « ga » et « wa ». Selon les auteurs (1973 : 92), la particule « wa » en japonais signifie que l'objet mentionné a été identifié antérieurement à l'énonciation des phrases et donc cet objet n'est pas considéré comme arbitraire.

²⁶ Choi J. J. (2016) distingue quatre types de phrases qui peuvent avoir différentes relations informationnelles : catégoriq ue, thétiq ue, définitoire et clivée. D'après l'auteur, les particules « -i / -ga » et « -eun / -neun » participent à la détermination du SN dont elles font partie, mais tout cela dépend aussi du caractère informationnel de la phrase.

Traduction en français : Le cheval court. / Un cheval court.

(6) 말이 달려온다.

mal+i dalryeo+o+n+da.

cheval+particule courir+venir+présent+terminaison.

Traduction en français : Le cheval vient (vers moi) en courant.

D'après ce qui est évoqué *supra* sur le rôle discursif de la particule, le SN « le cheval » de chaque exemple est spécifique : la particule « -i » est employée dans les deux cas montrant ainsi que ces chevaux sont thétiques. Toutefois, les deux occurrences de « cheval » possèdent différents degrés de saillance. Le « cheval » de (6) est spécifique, alors que l'exemple (5) peut être interprété soit comme « le cheval court », soit comme « un cheval court ». Ce point devient plus évident si nous remplaçons simplement la particule « -i » par « -eun » comme dans les exemples suivants :

(5')말은 달린다.

mal+eun dalri+n+da.

cheval+particule courir+présent+terminaison.

Traduction en français : Le cheval (générique) court.

(6') #말은 달려온다.

mal+eun dalryeo+o+n+da.²⁷

cheval+particule courir+venir+présent+terminaison.

Traduction en français : ? Le cheval (générique) vient (vers moi) en courant.

La phrase (5') ne pose aucun problème que ce soit d'un point de vue grammatical ou pragmatique. Néanmoins, dans cet exemple, « le cheval » est identifié de manière plus générique que celui de (5) en raison du changement de particule. Par ailleurs, l'exemple (6') semble peu naturel : un contexte particulier doit suivre pour qu'il soit plausible. En général, dans un tel cas, on essaierait d'interpréter la phrase en choisissant la fonction canonique de la particule auxiliaire qui est contrastive, c'est-à-dire qu'on présupposerait l'existence implicite d'un objet comparable dans le contexte ainsi que « le cheval vient vers moi en courant alors que le chien reste assis ».

Malgré le rôle important que joue la particule dans la détermination d'un nom, elle n'est

²⁷ Nous utilisons la signe # pour indiquer le fait que la phrase ne peut être qu'interprétée si un autre contexte ou une autre circonstance est supposé(e).

pas l'élément unique et décisif équivalant à l'article. Il se pourrait qu'un mot reste indéterminé quelle que soit la particule employée. La particule peut, pour une part, concourir à la détermination du nom, mais d'autres éléments entrent en ligne de compte et nous précisons ces éléments dans la partie qui suit.

2.2.2 Influence du prédicat : le sens et la forme de prédicat

En analysant les exemples ci-dessus, nous avons remarqué que la particule, en soi ne peut pas remplacer le système d'article du français. En coréen, le sens du prédicat, son aspect et son mode peuvent également avoir des répercussions sur la détermination du nom.

2.2.2.1 Le sens de prédicat

D'après Oh C. Y. (2012 : 83-85), la détermination dépend en substance de la manière dont on perçoit un événement et cela est souvent impliqué dans le verbe. Reprenons les exemples (5) et (6) :

(5) mal+i dalri+n+da.

cheval+particule courir+présent+terminaison.

Traduction en français : Le cheval court. / Un cheval court.

(6) mal+i dalryeo+o+n+da.

cheval+particule courir+venir+présent+terminaison.

Traduction en français : Le cheval vient (vers moi) en courant.

La différence entre ces deux exemples tient au prédicat. Celui de l'exemple (5), « dalri+n+da », est plus neutre que celui de (6), « dalryeo+o+n+da ». Ce dernier possède l'élément sémantique « venir » et comprend donc un trait directionnel ou d'orientation. D'ailleurs, ce verbe a un caractère éphémère, tandis que le verbe « dalri+n+da » de (5) désigne une action constante ou susceptible de se répéter. Tout cela rend « le cheval » de l'exemple (6) plus identifiable et plus spécifique que celui de (5) pour le locuteur et son interlocuteur. Cette différence résultant des caractères différents des prédicats provoque également la différence d'interprétation des phrases (5') et (6') où la particule « -i » est remplacée par « -eun » :

(5') mal+eun dalri+n+da.

Cheval+particule courir+présent+terminaison.

Traduction en français : Le cheval (générique) court.

(6') # mal+eun dalryeo+o+n+da.

Cheval+particule courir+venir+présent+terminaison.

Traduction en français : ? Le cheval (générique) vient (vers moi) en courant.

Dans ces exemples, le nom « mal » (le cheval), référent générique en raison de l'emploi de la particule « -eun », ne convient pas à un prédicat comprenant le trait spécifiant la direction de l'action comme l'est « dalryeo+o+n+da » de (6') : cette sorte de prédicat nécessite que « le cheval » soit saillant, alors que le caractère du SN « mal+eun » de (6') reste générique.

2.2.2.2 La forme de prédicat : choix de terminaison

La forme d'un prédicat concerne plusieurs notions linguistiques comme le *temps*, l'*aspect*, et le *mode*.

D'après la *GMF*, l'*aspect* est décrit ainsi :

[...] d'un point de vue interne, le procès peut être envisagé en lui-même, 'sous l'angle de son déroulement interne' (P. Imbs 1960). En effet, indépendamment de toutes considérations chronologiques, un procès dynamique implique en lui-même du temps, une durée plus ou moins longue pour se dérouler et se réaliser. On peut concevoir ce déroulement interne de façon globale ou l'analyser dans ses phases successives (de son début à sa fin). (Riegel *et al.*, 2018 : 517-518)

Traditionnellement, dans la grammaire coréenne, la notion d'*aspect* n'a pas été finement distinguée de celle de *temps* : le coréen est une langue à dominante temporelle. En d'autres termes, contrairement aux langues aspectuelles qui marquent nécessairement l'*aspect* dans la phrase, le coréen n'a pas de formule claire pour exprimer celui-ci (Nam K. S. *et al.*, 2019 : 397). L'*aspect* qui concerne le temps interne du verbe est montré par le niveau lexical, ainsi que par la flexion du verbe et, en coréen, la notion d'*aspect* risque de se mélanger avec celle du *temps* à cause de la flexion du verbe. Oh C. Y. (2012 : 86) souligne que l'*aspect* perfectif est exprimé par « -eot » dans les verbes qui correspondent à la forme du temps passé : ce type de verbe peut indiquer que l'événement est achevé. Autrement dit, l'événement est déjà réalisé. Reprenons l'exemple (2) mentionné *supra* :

(2) gangaji+neun sagwa+leul meog+eot+da.

chiot+cas-aux pomme+cas-objet manger+passé+terminaison

Traduction en français : Le chiot a mangé la pomme. (mais pas le chat)

Comme le prédicat « a mangé (meog+eot+da) » a une forme verbale perfective, l'acte de « manger » est réalisé, et par conséquent, les actants de ce prédicat, « chiot » et « pomme », sont concrets et référentiels. Si nous remplaçons le prédicat de (2) par celui de (2'), « mange (meog+neun+da) », qui est au temps présent et à l'aspect imperfectif, les noms « chiot » et « pomme » ne seront pas concrets²⁸ : cette phrase est une proposition générique. Plus spécifiquement, il s'agit d'une question générale sur l'alimentation du chiot :

(2') gangaji+neun sagwa+leul meog+neun+da.

chiot+cas-aux pomme+cas-objet manger+présent+terminaison

Traduction en français : (en général) Le chiot mange la pomme.

Ensuite, Oh C. Y. (2012 : 85-95) insiste sur le fait que les modes verbaux qui reflètent l'attitude du locuteur à l'égard de son énoncé sont présentés *via* la *terminaison pré-finale* ou celle des prédicats²⁹ en coréen. Pourtant, dans son étude, l'auteur ne distingue pas exactement la notion de *mode* de celle de *modalité*. C'est compréhensible, considérant le fait que ces notions sont empruntées à la linguistique occidentale et appliquées au coréen qui est totalement différente des langues occidentales : les recherches sur ces notions sont débattues, mais ce sujet reste encore flou dans la linguistique coréenne. Dans cette étude, nous n'entrerons pas dans le détail des différences de définitions entre les linguistiques coréenne vs. occidentale. Nous nous épargnerons la définition des notions de *mode* et de *modalité* d'autant qu'elles ne concernent pas immédiatement notre propos.

Par conséquent, il nous suffira de suivre les définitions proposées par la *GMF* sur les notions de *modalité* et de *mode*. D'après cet ouvrage, la modalité est définie ainsi :

Dans l'étude de la langue, les modalités sont considérées comme des éléments qui expriment un certain type d'attitude du locuteur par rapport à son énoncé. (Riegel *et al.*, 2018 : 975)

Par la suite, en français, on distingue cinq types de mode (l'indicatif, le subjonctif,

²⁸ Nous parlons ici de l'interprétation littéraire, en d'autres termes, nous ne considérons pas le contexte d'énonciation.

²⁹ Reprenons la phrase de l'exemple (1) : gangaji+ga sagwa+leul meog+eot+da. La terminaison pré-finale du prédicat « meogotda » est « -eot » et sa terminaison est « -da ».

l'impératif, l'infinitif et le participe), et la notion de *mode* est définie comme suit :

La définition traditionnelle des modes s'appuie sur la notion de *modalité* : les modes expriment l'attitude du sujet parlant à l'égard de son énoncé ; ils manifestent différentes manières d'envisager le procès. (Riegel *et al.*, 2018 : 511)

Les deux notions sont similaires dans la mesure où elles concernent la façon dont le locuteur perçoit l'événement. Cependant, la modalité peut marquer l'attitude du locuteur vis-à-vis du contenu de l'énoncé par différents moyens lexicaux ou syntaxiques, alors que le mode est le résultat de la grammaticalisation des modalités exprimées par les verbes. Dans cette optique, nous sommes essentiellement concernée par le mode et la façon dont il est représenté par les terminaisons finale et pré-finale, dans cette partie consacrée à la forme du prédicat.

Oh C. Y. (2012 : 85-95) souligne que le mode peut jouer un rôle important dans la détermination d'un nom. Nous remplaçons la terminaison du prédicat « meog+eot+da » de l'exemple (7) par « -nae », variante de la terminaison de simple description indiquant le présent. Puis, dans l'exemple (8), la terminaison pré-finale « -deo » est employée, ce qui signifie qu'il s'agit d'un événement du passé :

(7) 강아지가 사과를 먹네
gangaji+ga sagwa+leul meog+nae.
chiot+cas-sujet pomme+cas-objet manger+terminaison

Traduction en français : (Je vois que) Le chiot est en train de manger la pomme.

(8) 강아지가 사과를 먹더라
gangaji+ga sagwa+leul meog+deo+ra
chiot+cas-sujet pomme+cas-objet manger+(rappel)+terminaison

Traduction en français : (Je me rappelle que) Le chiot a mangé la pomme.

En choisissant la terminaison « -nae », le locuteur exprime qu'il est en train d'observer le chiot qui mange la pomme ou une pomme. En outre, par la terminaison pré-finale « -deo » de (8) qui est un marqueur d'évidentialité³⁰, nous pouvons savoir que la phrase décrit un événement dans le passé, vu ou perçu par le locuteur sans intermédiaire. Dans de tels cas,

³⁰ La notion d'*évidentialité*, selon Barbet & de Saussure (2012 : 3-5), concerne la source de l'information véhiculée par l'énoncé : l'information communiquée peut se baser directement sur la perception de locuteur ou indirectement sur la reprise d'autrui ou sur l'inférence.

les noms obtiennent la référentialité.

Ainsi, la particule et le prédicat employés dans la phrase peuvent, de façon diversifiée, contribuer à combler le manque d'articles en coréen : dans l'ensemble, ces éléments permettent de déterminer si l'événement qu'on désigne par le prédicat possède la substantialité, à savoir s'il est actualisé ou s'il reste un énoncé virtuel employé de manière générique.

2.3 La dépendance contextuelle ou situationnelle

La deuxième tendance du coréen à mentionner dans le domaine de CR est que cette langue dépend beaucoup du contexte ou de la situation de communication.

En synthétisant les fonctionnements de la particule « -i /-ga » et « -eun / -neun », Kim M. H. (2011 : 60) souligne que le coréen nécessite un processus cognitif subtil en raison de l'utilisation de particules pouvant être casuelles, thématiques ou connotatives selon la situation.

Nous avons également vu que la détermination en coréen dépend de manière synthétique de la particule et du prédicat qui font partie du contexte au niveau phrastique. En d'autres termes, il n'y a pas qu'un seul dispositif linguistique qui pallie l'absence de système d'articles. C'est en fonction de la situation contextuelle du segment linguistique, issu de tel ou tel choix, que les constituants d'une phrase obtiennent la détermination.

La haute dépendance contextuelle de la langue coréenne est également prouvée par le phénomène d'ellipse. Ahn J. K. & Eun H. Y. (2006) soulignent que le contexte affecte le fait que le coréen emploie fréquemment les pronoms zéros non seulement pour la position sujet, mais aussi pour les positions des autres constituants, y compris celle d'objet. Plus particulièrement, les pronoms personnels de dialogue et de la troisième personne en français ne sont pas souvent réalisés en coréen, notamment à l'oral. Regardons l'exemple suivant :

(9) A : (너) 개 어떻게 생각해 ?

(neo) gyae eotteokae saenggak+h+ae ?

(tu) elle/lui comment penser+terminaison-interrogative.

Traduction en français : Qu'est-ce que ø penses d'elle/de lui ?

(= Qu'est-ce que tu penses d'elle/de lui ?)

B : (나는) (개를) 좋아하지.
(na+neun) (gyae+leul) joaha+ji.
(je+marque thématique) (elle/lui+cas-objet) aimer+terminaison
Traduction en français : ∅ ∅ aime (= Je l'aime.)

Dans la question (A) de (9), le sujet « tu » (neo) est ellipsé. Il en va de même dans la réponse (B) pour « je » (na+neun) et « elle » (gyae+leul) qui sont respectivement le sujet et l'objet de l'énoncé. Pourtant, cette ellipse n'empêche pas la compréhension parce que, grâce au contexte d'énonciation, nous pouvons déduire le sujet et l'objet dont on parle.

Comme nous venons de le voir, la haute dépendance contextuelle a des répercussions sur plusieurs phénomènes linguistiques en coréen. De prime abord, le contexte phrastique fonctionne comme un des éléments comblant le manque d'article. Ensuite, le contexte, de manière plus générale, justifie le phénomène d'ellipse, à savoir celui de pronom zéro.

Bilan

Ce chapitre visait à rendre compte de notre point de vue sur le phénomène des CR, le point de vue contrastif inter-langues, et à fonder une base de compréhension de la langue coréenne qui fait l'objet de la comparaison.

Les études antérieures des CR d'un point de vue contrastif inter-langues signalent que malgré les différences typologiques entre les langues, certaines similarités dues à la particularité du genre textuel peuvent être relevées (voir la section 1.1). En d'autres mots, comprendre non seulement les disparités des langues, mais aussi la structuration du texte elle-même, est importante dans une perspective comparative, celle-ci étant l'un des facteurs qui influence la caractérisation des CR.

Par ailleurs, nous avons vu que les CR sont des phénomènes linguistiques peu introduits dans la linguistique coréenne (voir la section 1.2). Faute d'études sur la notion, notamment d'un point de vue inter-langues, cette étude est un projet novateur pour ceux qui souhaitent travailler sur les CR en Corée.

Dans la deuxième partie de ce chapitre, nous avons présenté les caractéristiques de la

langue coréenne qui pourraient agir sur les CR : nous nous concentrons particulièrement sur le fait que le coréen n'a pas de système d'article. Pour cette raison, cette langue possède un système quasi-alternatif pour combler le manque. La réalisation de la détermination d'un référent en coréen est faite, de manière indirecte, par les éléments de la phrase, la particule et le prédicat, ainsi que par le contexte d'énonciation.

Chapitre 4.

Les expressions référentielles en français et en coréen

Introduction

Nous avons travaillé sur les notions-clés et les caractéristiques essentielles des CR dans les deux premiers chapitres de notre recherche. Ensuite, pour circonscrire le domaine de notre étude, le chapitre précédent a traité, d'une part, le point de vue contrastif appliqué aux expressions référentielles et aux CR par nos prédécesseurs et, d'autre part, les particularités de la langue coréenne à noter, avant de mener une recherche inter-langues.

Toutefois, il reste encore une question fondamentale pour laquelle nous n'avons pas encore trouvé de réponse consensuelle : à quoi correspond exactement la notion d'*expression référentielle* ? Autrement dit, y a-t-il une définition claire et précise délimitant celle-ci ?

Afin d'y apporter une réponse, ce chapitre fixe le double objectif d'explicitier les points suivants : la définition et la délimitation d'*expression référentielle* (§1) et les caractéristiques de ses sous-catégories, les noms propres (§2), les pronoms (§3) et les SN anaphoriques (§4) dans les deux langues.

Ainsi, nous espérons pouvoir poser le cadre théorique de cette thèse qui vise, en fin de compte, à construire les conventions d'annotation des CR susceptibles de s'appliquer aux textes français et coréen (chapitre 6) et à analyser les différences inter-langues présentées *via* le phénomène des CR (chapitre 7).

1 Le classement des expressions référentielles

Nous avons succinctement discuté la présence des différentes conceptions d'*expression référentielle* dans les recherches antérieures portées sur les CR (voir la section 2.1 du chapitre 2). Le présent chapitre, qui vise à décrire les expressions référentielles du français et du coréen, s'attachera à une discussion générale sur leurs caractéristiques. Avant d'entamer l'explication de chaque catégorie grammaticale, dans cette sous-section, nous nous proposons de faire un bilan des différentes approches sur la définition et la délimitation des expressions référentielles. En rappelant ceux qui sont traités auparavant, nous envisageons de préciser notre position sur le sujet.

1.1 Les expressions référentielles et le type de référence

Le choix des expressions référentielles varie en fonction de la situation de référence. La *GMF* (2018 : 965-967) distingue la *référence* en quatre sous-catégories et propose des différents types d'expression qui s'emploient au sein d'une référence indexicale, une référence descriptive, une référence dénominative et une référence pronominale.

En premier lieu, la référence indexicale a un caractère déictique, c'est-à-dire que le sens codé de l'expression concerne immédiatement un participant ou une composante de l'énonciation (par exemple, *Maintenant, je m'installe ici.*). Ensuite, la référence descriptive concerne une expression nominale (définie, possessive et/ou démonstrative) susceptible de décrire des caractéristiques du référent (ex. *Le président de la République a commencé ses vacances*). En revanche, la référence dénominative identifie et désigne le référent sans ambiguïté *via* l'utilisation du nom propre (ex. *Emmanuel Macron a commencé ses vacances*). Enfin, la référence pronominale emploie des expressions qui n'ont pas de contenu descriptif en elles-mêmes ou qui ont un faible contenu descriptif (ex. *J'adore ça ! / Quelqu'un t'a cherché*).

Cette distinction générale sur les expressions référentielles appuie l'idée que certaines d'entre elles sont référentiellement autonomes alors que d'autres dépendent du contexte ou de la situation d'énonciation pour être interprétées. Par exemple, *Emmanuel Macron* renvoie toujours à un même référent quel que soit le contexte tandis que l'interprétation du SN *le président de la République* ou du pronom personnel *il* peut varier selon le contexte d'énonciation.

1.2 Les expressions référentielles dans le cadre de la chaîne de référence

Hormis la conception générale proposée par la *GMF* qui vient d'être mentionnée, plusieurs auteurs proposent des approches diversifiées sur les expressions référentielles sous l'angle de leur capacité d'être employée comme maillon. Nous introduisons brièvement quelques approches concernant le champ de notre étude.

Ariel (1990) classe les expressions référentielles selon les différents degrés de l'accessibilité référentielle¹. De ce point de vue « relativement ouvert », nous pouvons inclure toute forme de SN comprenant l'ellipse ou l'accord verbal qui dépendent tous deux du contexte pour être interprétés.

Charolles (2002) distingue, pour sa part, les cinq types d'expressions référentielles les plus usitées en français (les noms propres, les SN définis, démonstratifs, indéfinis et les pronoms) et décrit leurs caractéristiques². Cependant cette classification n'est pas exhaustive et l'auteur reconnaît également la nécessité d'ajouter d'autres expressions telles que les possessifs et les SN sans déterminants.

La typologie proposée par Baumer (2017) semble à la fois plus simple et plus étendue. Dans le cadre de l'analyse des CR, il définit trois types d'expressions référentielles : les pronoms personnels anaphoriques, les noms propres et les syntagmes nominaux coréférentiels ou anaphoriques renvoyant aux « trois grandes catégories de marqueurs référentiels indexicaux constituant les maillons des CR » (*ibid* : 73).

Dans la linguistique coréenne, Chang K. H. (2004 : 54) analyse les expressions référentielles coréennes et met en avant que leur classement basé sur un critère morphosyntaxique ne suffit pas à rendre compte de la nature et du fonctionnement de la référence : elle propose donc de les classer en deux catégories, une substantive et l'autre démonstrative³. D'après cette classification, le substantif renvoie non seulement au référent, mais aussi à une information essentielle. En ce qui concerne les démonstratifs, ceux-ci

¹ Voir la section 2.1 (notamment la figure 4) du chapitre 2.

² Également la section 2.1 du chapitre 2.

³ Il s'agit de trois expressions de *jisi gwanhyungsa* (déterminant démonstratif), les expressions ostensives basées sur la distance. Comme « this / that / it » en anglais, le coréen distingue les trois dimensions de démonstratifs selon la distance de l'objet, « i / geu / jeo ». Pourtant, ces derniers se distinguent de ceux de l'anglais par le pivot de la référence : soit qu'il est le locuteur soit qu'il est l'interlocuteur. Nous aurons l'occasion d'en reparler dans le chapitre 6.

indiquent la localisation et ne concernent pas les propriétés du référent.

Toutes ces approches, aussi différentes soient-elles, ne sont pas si éloignées les unes des autres dans le sens qu'elles visent à distinguer les expressions référentielles en fonction de leur propriété référentielle. Il est néanmoins important de déterminer le classement le plus adapté à notre recherche concernant les CR : notre choix se doit à la fois de prendre en compte les critères qui reflètent la fréquence d'occurrence des catégories grammaticales liées à notre travail d'annotation⁴ et d'être des plus pertinent pour la comparaison des deux langues complètement différentes.

Dans cette perspective, nous suivrons les conceptions de Baumer (2017) et Charolles (2022) pour décrire les expressions référentielles dans les deux langues leurs idées étant plus adaptées à notre propos. Un classement des noms propres, pronoms personnels et des syntagmes nominaux (sous-classés encore en définis, indéfinis et démonstratifs chez Charolles) en trois types pourrait paraître moins détaillé par rapport à l'approche d'Ariel ou d'autres conceptions plus ouvertes comme celles évoquées en amont dans notre étude (voir la section 2.1 du chapitre 2). Cependant, ce choix « relativement lâche » convient à notre propos parce qu'il facilite la comparaison de ces deux langues ayant un système langagier complètement différent. Dans la mesure où les articles n'existent pas en coréen, il serait inadéquat de comparer directement et immédiatement les expressions référentielles du coréen à celles du français dans lesquelles l'article joue un rôle important. Par conséquent, en suivant ces deux auteurs, nous essaierons de rendre compte des caractéristiques et des emplois de chaque catégorie grammaticale dans les deux langues afin de montrer la manière dont le coréen constitue les SN. Inexistantes en coréen, celles qui ne sont pas traitées dans ce chapitre, le déterminant possessif ou le pronom relatif, seront mentionnées dans le chapitre 6 avec leur mode de traitement dans l'annotation.

2 Les noms propres

Les noms propres (désormais Npr) présentent quelques caractéristiques qui les

⁴ L'annotation se porte sur notre corpus constitué d'un roman français et d'un roman coréen. On aura l'occasion d'en parler dans le chapitre 5.

différencient des noms communs (désormais Nco). Ils peuvent référer à des personnes, des objets, des lieux (y compris des villes, des montagnes, des pays, *etc.*), des monuments, des institutions, *etc.* comme les Nco. Cependant, les Npr se distinguent des Nco par le lien qu'ils établissent avec les objets du monde.

Une ville se caractérise par son nom autant sinon plus que par son statut politique (préfecture, capitale ou simple municipalité), le nombre de ses habitants ou sa position géographique. (Gary-Prieur & Noailly, 2019 : 1)

Ainsi, contrairement aux autres éléments de la langue, les Npr ont un double statut : l'objet de langue et l'objet du monde. Comme expliqué en amont, cette particularité est liée à la raison pour laquelle Saussure et les structuralistes ont souvent été amenés à contourner le problème de la référence⁵.

Dans les pages suivantes, nous détaillerons la notion de *Npr* et la manière dont elle a suscité un regain d'intérêt chez les linguistes. Puis nous rappellerons leurs caractéristiques et fonctions d'un point de vue référentiel. Pour ce faire, nous nous inspirerons principalement des linguistes, Jonasson (1994), Charolles (2002) ainsi que de Gary-Prieur & Noailly (2019), qui ont proposé la synthèse la plus récente sur le sujet.

2.1 La Signification de nom propre

Depuis longtemps, la question de la nature de Npr est au centre des préoccupations de nombreux philosophes. Charolles (2002) présente deux approches ayant trait au problème de la signification de Npr : la première défend l'idée que les Npr ne possèdent pas de sens alors que l'autre insiste sur le fait qu'ils ont un sens et un caractère descriptif.

Selon Mill (1724/1988) et Kripke (1972/1982) (cités par Charolles, 2002 : 54-55 et Gary-Prieur & Noailly, 2019 : 2), les Npr dénotent mais ne connotent pas, c'est-à-dire qu'ils ont une fonction dénotative mais n'évoquent aucun des attributs de leur référent. Pour identifier celui-ci, le destinataire doit de connaître préalablement ce qui est nommé par ce Npr, à savoir un lien conventionnel entre le Npr et le référent établi par l'*acte de baptême* et n'ayant aucune relation avec la nature du référent. Ainsi, Kripke souligne que les Npr constituent des « désignateurs rigides » :

⁵ Voir la section 2.1 du chapitre 1.

Le Np⁶ est donc pour Kripke un désignateur rigide, c'est-à-dire qu'il désigne le même objet dans tous les mondes possibles associés à un énoncé. (Gary-Prieur & Noailly, 2019 : 2)

Les Np étant liés à leur porteur par une convention qui est indépendante de leurs caractéristiques substantielles peuvent continuer à référer à ceux-ci quels que soient les changements susceptibles de les affecter. (Charolles 2002 : 55)

Considérant que le Npr désigne toujours le même particulier et que ce lien est stocké dans la mémoire à long terme, le Npr ne doit pas être « senti comme une description de ce particulier, vu que celui-ci peut changer, être différent d'un moment à l'autre sans pour cela perdre son identité fondamentale que le Npr doit saisir » (Jonasson, 1994 : 17). Charolles décrit également ce caractère du Npr en le comparant avec la description définie :

Les Np désignent un particulier dans tous ses états, dans toutes les phases de son existence, sous toutes ses apparences, dans ce qui constitue son essence individuelle, alors que les descriptions définies n'exploitent jamais que des traits accidentels. (Charolles 2002 : 55)

Pour reprendre l'exemple mentionné *supra*, les deux expressions, « Emmanuel Macron » et « l'actuel Président de la République Française », qui renvoient au même référent dans le contexte politique de 2023, peuvent cependant avoir des effets pragma-sémantiques différents. La description définie « l'actuel Président de la République Française » peut être efficace lorsque nous nous adressons à des écoliers coréens qui ne sont pas familiers avec les politiciens français. Cependant, si le moment de l'énonciation a eu lieu en 2006, le référent de cette description définie change. En effet, elle renvoie à une autre personne qui a exercé la fonction de président à cette date, alors que le référent spécifique du Npr « Emmanuel Macron » ne change pas : les Npr, démunis d'une signification linguistique, restent indifférents aux traits fluctuants des référents. Pour ainsi dire, « Emmanuel Macron » est pour toujours « Emmanuel Macron ». Sous cet aspect, « la désignation par Np, en même temps qu'elle est 'rigide' et permanente, est parfaitement docile à une représentation évolutive du référent » (Gary-Prieur & Noailly, 2019 : 22) parce que les Npr sont liés directement à des êtres particuliers.

Cette idée sur la rigidité des Npr est contestée par plusieurs auteurs comme Frege (1879/1971), Russell (1905), Strawson (1959/1973) et Searle (1972) (cités par Charolles,

⁶ Pour la citation, nous reprenons l'abréviation utilisée par l'auteur.

2002 : 55-57). Ils s'accordent sur l'idée que tout Npr a un sens même si leurs arguments diffèrent dans le détail.

(1) a. R. Gary est E. Ajar.

b. L'Everest est le Chomolungma.

Notre compréhension des phrases équatives de l'exemple (1) (phrases que nous reprenons à Charolles, 2002 : 56) et la façon dont nous jugeons la valeur de vérité d'un énoncé proviennent du fait que les Npr possèdent un sens. Dans le cas contraire, la phrase de (1a) semblerait tautologique. Selon Searle (1972)⁷, les Npr fonctionnent, en réalité, comme des outils commodes qui permettent au locuteur de se référer à des êtres particuliers. En d'autres termes, lorsqu'un Npr est employé, il n'existe pas de possibilité pour qu'il soit interprété de manière ambiguë⁸. Cependant, il n'en demeure pas moins qu'entre les Npr et les être particuliers, il y a « une relation descriptive (substantielle) quelconque » (Charolles, 2002 : 57).

Kleiber (1981) partage le point de vue ci-dessus mais celui-ci est plus nuancé. Il propose la notion de *prédicat de dénomination* qui stipule que les Npr seraient la forme abrégée de la glose « être appelé Npr ». Dans cette optique, à l'aide de la forme « le X appelé R. Gary » qui exclut tout autre référent, nous pouvons constituer le sens du Npr « R. Gary ». L'auteur abandonnera ultérieurement cette idée (Kleiber, 1996 : 573). Cependant, il ne renonce pas à l'idée que les Npr ont un sens dénominatif⁹. Il insiste sur le fait que la relation unique entre le Npr et son porteur, autrement dit le caractère dénominatif du Npr, est codé dans son usage.

Ces différentes conceptions sont cependant complémentaires pour Wilmet (1991 : 114-115) qui souligne que le sens du *Npr* peut être différencié selon la façon dont on circonscrit le plan linguistique. Lorsqu'un Npr est employé sur le plan de la langue, c'est un « asémantème », alors que, dans un discours, il possède une valeur descriptive : le contexte et la situation d'énonciation peuvent influencer la fonctionnalité sémantique des Npr.

Bien que le point de vue sur la signification des Npr diffère selon les auteurs, le consensus porte sur le fait que ces formes possèdent une fonction référentielle. Dans cette étude, nous

⁷ Également cités par Charolles (2002 : 57).

⁸ Il peut y avoir des cas d'homonymie comme le cas du nom « Pierre Dupont » ou bien celui du nom de la ville « Paris » (France et Texas). Pour l'instant, nous ne tenons pas compte de ces cas particuliers.

⁹ Cf. la section 2.3.2.1 *infra*.

adaptions l'idée que les Npr eux-mêmes sont dénués de sens, mais qu'ils construisent leur sens à partir du lien qu'ils entretiennent avec le contexte et la situation d'énonciation. Partant de cette perspective, nous examinerons les caractéristiques du Npr, notamment sa fonction référentielle, dans les pages suivantes.

2.2 Caractéristiques et fonctions référentielles du nom propre

Charolles (2002 : 73-74) synthétise les caractéristiques des Npr comme suit :

[...] les noms propres occupent une place à part parmi les expressions référentielles : ils désignent des êtres particuliers en les distinguant au sein de leur catégorie d'appartenance sans faire allusion par eux-mêmes à cette catégorie. Le lien qui unit le nom propre à son porteur n'est pas fondé sur une relation descriptive, mais sur une convention dénomminative, ce qui ne l'empêche pas de référer à celui-ci (même dans les emplois indirects) et non au nom qui sert à le désigner. Les noms propres se chargent de sens encyclopédique et peuvent, de ce fait, finir par signifier des attributs notoirement affectés aux individus qu'ils désignent. Enfin, les noms propres réfèrent à des particuliers qui ne sont pas supposés déjà présents dans l'esprit des interlocuteurs : ils attirent l'attention sur ces particuliers en vue de renvois ultérieurs. (Charolles, 2002 : 73-74)

Comme l'explique l'auteur, les Npr renvoient aux êtres particuliers en les distinguant des autres de la même catégorie d'appartenance et suscitent l'attention des interlocuteurs sur ces individus concernés. Jonasson (1994) souligne également que la désignation d'un particulier est réalisée *via* la fonction cognitive des Npr, à savoir « la fonction de nommer et de maintenir une individualité, de fixer une identité » (*ibid* : 18). Mais, pour que cette référence soit réussie, comme nous l'avons mentionné *supra*, il est nécessaire de procéder à un premier moment de dénomination souvent appelé le « baptême linguistique » :

[...] pour qu'un nom propre identifie valablement un individu particulier, il faut qu'il lui ait été préalablement assigné par un acte *ad hoc* de 'baptême linguistique' - parfois perdu dans la nuit des temps – en tant qu'occurrence particulière d'une catégorie nominale... (Riegel *et al.*, 2018 : 336)

Le référent ainsi nommé est définitivement identifié par ce Npr : le porteur du Npr est perçu et reconnu directement dans la mémoire des interlocuteurs. Toutefois, il convient de souligner que l'exigence du « baptême linguistique » peut être ignorée dans certains cas. Voici

une situation présentée par Charolles (2002 : 67). Lorsque nous lisons un roman écrit dans une langue étrangère, il arrive que les Npr employés afin de désigner des personnages ou des lieux soient très différents de ceux de notre propre langue. Cependant, les lecteurs classent sans aucun problème ces Npr correctement dans leurs catégories d'appartenance :

Le Np ne délivre par lui-même aucune des informations descriptives faisant que tel particulier est particulier, mais disant qu'il est particulier, son interprétation doit au minimum aboutir à la compréhension de la classe (être humain, bateau, cheval, etc.) au sein de laquelle il est particulier du fait qu'il est appelé Untel. (Charolles, 2002 : 67)

C'est ce que Gary-Prieur & Noailly (2019 : 22) appellent l'« immersion immédiate » : il se peut que dans les débuts de romans, l'auteur présente les référents comme s'ils « étaient déjà connus du lecteur ». Ensuite, le lecteur collecte les informations utiles à partir des autres éléments du texte et parvient à placer le référent dans sa catégorie d'appartenance. Ce processus peut s'appuyer sur les dispositifs linguistiques que sont l'apposition, le prédicat, ou bien les caractéristiques du genre textuel. Regardons l'exemple suivant qui est un extrait d'un roman :

(2) **Melquiades**, qui était un homme honnête, le mit en garde : « ça ne sert pas à ça ». Mais JOSE ARCADIO BUENDIA, en ce temps-là, ne croyait pas à l'honnêteté des gitans, et il troqua son mulet et un troupeau de chèvres contre les deux lingots aimantés. *Ursula Iguaran*, sa femme, qui comptait sur ces animaux pour agrandir le patrimoine domestique en régression, ne parvint pas à l'en dissuader. — (extrait du roman *Cent ans de solitude* par Gabriel García Márquez, p.18).

Cet extrait fait usage de trois Npr (non français), *Melquiades*, *Jose Arcadio Buendia* et *Ursula Iguaran*, pour désigner des personnages. Leur introduction suggère que des référents porteurs de tels noms existent. Cependant, les informations sur ces référents se cumulent au fur et à mesure de la progression du texte. D'après le segment apposé et la phrase subséquente, « Melquiades » est un gitan honnête, « José Arcadio Buendia » un homme sédentaire qui a un mulet et un troupeau de chèvres et « Ursula Iguaran » est la femme de José. Ainsi, au fil de la lecture et en vertu de certaines inférences, les lecteurs construisent le « sens encyclopédique » des Npr (ou le sens fictif fourni par le texte quand il s'agit de fiction), c'est-à-dire tout ce qu'il faut savoir sur le référent de ces Npr.

Ainsi, lorsqu'un particulier est désigné par un Npr et lorsque sa classe est identifiée, la référence est effective et les Npr servent de « repère existentiel ». Autrement dit, leur emploi

signale que les porteurs de ces termes existent. Dès lors, « l'auditeur et le lecteur leur rattachent toutes sortes d'informations qui vont, non pas enrichir leur signification, mais compléter la connaissance qu'ils peuvent avoir de leur porteur » (Charolles, 2002 : 68). Ce processus de référence est schématisé comme suit par Wilmet (1991 : 115) :

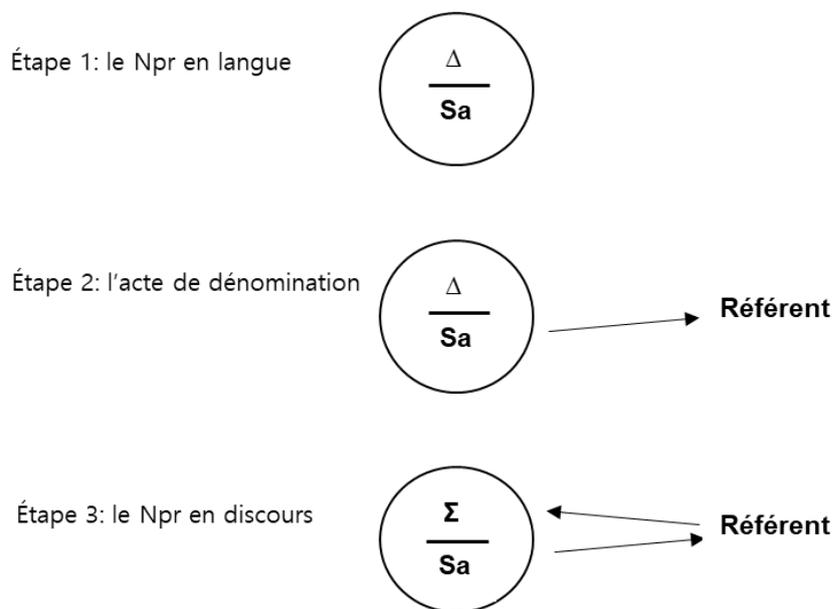


Figure 5 Processus de la référence des Npr (Wilmet, 1991 : 115)

D'après l'auteur, les Npr sont primordialement des signes munis de signifiant (Sa) et de signifié vide (Δ). Donc, un Npr peut renvoyer à plusieurs référents potentiels à cette étape (Par exemple, le Npr « Tintin » peut nommer un homme, un animal de compagnie, un endroit, *etc.*). Lorsque le Npr est connecté avec un référent spécifique parmi les référents potentiels, le noyau atomique de signification commence à être rempli. Puis, en discours, ce Npr collecte la somme des informations empiriquement accumulées, ce qui est exprimé par « sigma (Σ) » dans ce schéma.

En raison de l'absence d'un sens présupposé, les Npr attirent l'attention des interlocuteurs : la construction du sens des Npr commence au moment où la référence devient effective. C'est à ce moment-là que le référent prend une place dans l'esprit des interlocuteurs. Cette fonction référentielle des Npr peut « servir à construire un objet de référence, c'est-à-dire introduire un nouveau référent dans le discours, ou bien reprendre un objet déjà construit » (Jonasson, 1994 : 65). De la même manière, Schnedecker (1997 : 51-52) souligne que, par la présence d'un Npr dans le texte, nous comprenons que le référent dénommé ainsi aura de l'importance et qu'il sera sans doute anaphorisé ultérieurement.

2.3 Les différentes formes des noms propres et leurs emplois

Nous venons de voir que, comme le souligne Wilmet (1991 : 115), les Npr sont étudiés suivant deux perspectives sur le plan sémantique : l'une au niveau de la langue et l'autre au discours. Dans la perspective de la langue, ils occupent une place moins intéressante parce que, sans contexte, ils ont « une intension nulle » (*ibid.*). Cependant, au sein du discours, nous pouvons facilement rencontrer divers types d'emplois des Npr.

Dans cette partie, nous distinguons, en premier lieu, les différentes réalisations formelles des Npr : les Npr nus (non modifiés) et ceux qui sont précédés d'un déterminant¹⁰. Les premiers occupent une place privilégiée dans de nombreuses recherches tandis que les derniers ont été plus tardivement étudiés comme le mentionne Leroy (2005) :

La description et l'analyse par les linguistes des constructions effectives du Npr dans les énoncés et dans les textes ont fait apparaître toute une variété d'emplois qui, bien que grammaticaux, étaient ignorés ou tenus pour marginaux dans les approches antérieures du Npr. (Leroy, 2005 : 3)

A partir de la théorie de Gardiner (1954, cité par Jonasson, 1994 : 174-181) classant ces deux types d'emplois sous la dénomination « Npr incarnés / Npr désincarnés » et celle de Boër (1978, cité *ibid.*), « Npr référentiels / Npr attributifs », nous commençons à élargir le regard sur la définition des Npr et leurs divers usages dans les discours.

Pour ce faire, nous nous concentrerons sur ces différents types de Npr et leurs emplois afin de mieux cerner la notion de *Npr* et nous tenterons d'appliquer cette dernière à notre travail. En dernier lieu, nous déterminerons également si la façon dont on définit les Npr en coréen est similaire à celle du français ; si tel est le cas, l'enjeu sera de caractériser les particularités des emplois des Npr dans une langue dépourvue d'articles.

2.3.1 Les noms propres non modifiés et l'emploi prototypique

Commençons par aborder la question des Npr nus (non modifiés). D'après la *GMF* (2018 :

¹⁰ De nombreuses recherches démontrent qu'il y a différents types de Npr selon leurs traits morphologiques : les Npr « purs » qui sont les Npr prototypiques (*Strasbourg, Kléber, etc.*), les Npr à base descriptive ou mixte (*le Mont Blanc, le Tour de France, etc.*) et les Npr possédant l'article défini de manière inhérente (*La Rochelle, Leblanc, etc.*). Cependant, ce classement basé sur les traits morphologiques des Npr ne fait pas l'objet principal de notre étude. Donc, nous n'entrons pas dans plus de détail.

337-338), à part certains cas où l'article défini fait partie du Npr (*La Rochelle, Le Havre etc.*), les Npr sont prototypiquement dépourvus de déterminant et de modifieur. Comme nous l'avons déjà mentionné (voir la section 2.2), ils sont autodéfinis renvoyant directement à leur référent et ont une fonction grammaticale équivalente aux syntagmes nominaux :

(3) a. **Lara** est venue chez moi.

b. **L'andorrane** est venue chez moi.

Le prénom « Lara » de (3a) signale l'existence d'une personne qui s'appelle ainsi et désigne directement le porteur de ce nom. Par rapport au SN de (3b), ce Npr est autonome aux niveaux de la détermination et de la référence. En revanche, le SN « l'andorrane » indique que la personne dont on parle est une femme et qui est de nationalité andorrane. Compte tenu de l'article « l' », cette expression désigne une personne connue ou, du moins, mentionnée préalablement par les interlocuteurs. Même si la façon dont on détermine le référent diffère, les Npr et les SN jouent la même fonction grammaticale dans la phrase.

Jonasson (1994 : 64-72) fait le départ entre trois emplois prototypiques des Npr de ce type : la fonction référentielle, la fonction vocative et celle de nomination. Les exemples suivants illustrent respectivement les trois cas :

(4) Tous les épisodes de **Lupin** sont maintenant disponibles sur **Netflix**.

(5) Ecoute-moi, **Emmanuelle** !

(6) a. Son nom de famille est **Mo**.

b. Son chat s'appelle **Pori**.

L'exemple (4) montre une fonction référentielle directe des Npr. Le Npr « Lupin » renvoie à un personnage de fiction et « Netflix » réfère à une plateforme de streaming américaine. Ensuite, les Npr en fonction vocative, comme dans l'exemple (5), sont utilisés afin d'interpeller le destinataire. Selon Jonasson (1994 : 67), dans ces emplois, les Npr ne sont pas syntaxiquement liés aux autres éléments de la phrase mais signalent plutôt une liaison entre l'énoncé et son destinataire. Finalement, la fonction nominative des Npr exprime « le lien dénominatif existant entre le Npr et une entité particulière » (Jonasson, 1994 : 69) comme dans (6). Ces emplois correspondent à un prototype, à l'usage dit « standard » des Npr. Cependant, lorsqu'ils sont modifiés par d'autres éléments linguistiques, l'analyse se complique, comme nous ne le découvrirons ci-dessous.

2.3.2 Les noms propres modifiés et les emplois non-prototypiques

Grâce à l'intérêt croissant des linguistes pour les usages réels des Npr dans les textes, il est admis que ceux-ci peuvent être employés de manière variée avec un déterminant ou des modifieurs dans l'objectif d'effectuer des effets linguistiques dans les textes comme le synthétise Leroy (2005) :

[...] En effet, celui-ci (le Npr), loin de se limiter à des constructions canoniques du type de *Paul danse* et à une fonction de désignation directe et immédiate d'un référent unique, se construit aisément et fréquemment avec un prédéterminant, singulier ou pluriel et/ou avec tel(s) ou tel(s) complément(s) ; il abandonne tout aussi facilement la position référentielle pour se trouver en position attributive ou en apposition ; enfin, il rompt ou modifie parfois les liens qui l'unissent à son référent attiré, pour ne le désigner que partiellement ou de façon 'oblique', voire pour renvoyer à un tout autre référent. (Leroy, 2005 : 3)

De la même manière, Kleiber (1981 : 332) définit un Npr modifié comme étant « accompagné de déterminants qui lui font perdre le caractère 'unique' ou 'singulier' fréquemment assimilé à la marque spécifique qui l'oppose aux noms communs ». L'emploi d'un tel Npr est donc considéré comme non-prototypique et, dans ce cadre, les Npr sont susceptibles d'être compris dans toutes les structures possibles du SN.

Ainsi, dans cette sous-partie, nous présentons les divers usages non-prototypiques des Npr afin de proposer une vue d'ensemble sur les Npr modifiés et de faciliter le processus d'annotation et l'analyse des CR. Il existe plusieurs façons de classer les utilisations des Npr. Toutefois, nous retiendrons le classement proposé par Jonasson (1994), à savoir celui concernant les emplois dénominatif et métaphorique. Nous compléterons cette classification par deux autres types souvent mentionnés par de nombreux auteurs comme Kleiber (1996), Gary-Prieur & Noailly (2019) : l'emploi métonymique et celui dit de fractionnement.

2.3.2.1 Type dénominatif

Il convient de rappeler ici la notion de *prédicat de dénomination* (voir la section 2.1.) proposée par Kleiber (1981) : « être appelé Npr ». Cette formule décrit les usages dits 'dénominatifs' qui concernent essentiellement « les noms de personnes, mais aussi les noms de ville, bourgades, villages » d'après Gary-Prieur & Noailly (2019 : 43). Voici la description de la notion présentée dans la *GMF* :

En emploi dénominatif, le nom propre renvoie à la classe de ceux qui le portent : *Un Meyer est venu me voir – Il n’y a pas d’Huguette au numéro que vous demandez – Aujourd’hui c’est la fête des Alfred*. Cet emploi peut être souligné par des marqueurs comme *certain, nommé* (*un certain Paul, un nommé La Rocca*) et implique généralement le défaut d’une connaissance plus précise du référent.

L’adjectif épithète peut qualifier un nom propre en emploi dénominatif précédé d’un article (*le grand Charles, la petite Fadette, la belle Hélène, etc.*) pour l’affecter d’une propriété jugée caractéristique (par exemple de nature) et qui le distingue des autres porteurs du nom propre. (Riegel *et al.*, 2018 : 338-339)

Comme indiqué, ce « prédicat de dénomination » peut se réaliser assez librement : au singulier (*un Npr, un certain/dénommé Npr*), au pluriel sous une forme d’un SN indéfini (*des Npr*) ou défini (*les Npr*) ou avec d’autres déterminants¹¹. Regardons l’exemple suivant :

(7) a. Il y a **trois Kim** à Strasbourg

b. **Un Kim** est venu te voir.

c. Moi, je connais **la jeune Kim**.

Gary-Prieur & Noailly (2019 : 44) distinguent deux types d’emplois dénominatifs. Le premier concerne *un Npr* considéré comme « nom de série », quand plusieurs entités portent le même nom et sont susceptibles d’être décrites par la construction « les X appelés Npr ». C’est le cas d’exemple (7a) : le Npr « Kim » renvoie à un sujet pluriel parce que nous faisons référence à une situation dans laquelle il y a littéralement trois personnes à Strasbourg portant le nom « Kim »¹². Ensuite, dans (7b-c), nous désignons un des « Kim ». Cette désignation, « Un Kim », correspond au deuxième type d’emploi dénominatif où le locuteur introduit un individu. Dans un tel cas, le SN « Un Kim » de (7b) limite notre connaissance du référent au fait qu’il est un homme. À travers cet emploi, le locuteur sous-entend qu’il ne dispose pas d’assez d’information sur le référent pour le nommer de manière concrète ou qu’il ne s’intéresse pas à une identification concrète du référent. En revanche, « la jeune Kim » de (7c) réfère à une personne particulière parmi les Kim : l’adjectif épithète caractérise le référent en

¹¹ Sauf le défini singulier parce que *le Npr* employé seul renvoie nécessairement au référent lui-même et qu’il n’y a pas de sens dénominatif dans cet emploi.

¹² L’emploi pluriel ici se comprend comme une « pluralité interne » (Gary-Prieur et Noailly, 2019 : 43) parce que le référent est humain qui n’est pas reproductible. Il faut donc considérer qu’il s’agit d’une référence collective.

fonctionnant de manière descriptive. Par conséquent, outre le fait que le référent est un individu qui porte le nom « Kim », l'article et l'adjectif épithète précisent qu'il s'agit d'une femme que le locuteur et l'interlocuteur connaissent et qu'elle possède un trait distinctif, « jeune », par rapport aux autres « Kim ».

Ainsi, les Npr de ce type ne fonctionnent plus comme des formes totalement identifiantes mais plutôt comme des séquences descriptives et classifiantes, voire caractérisantes. Dans cet emploi, l'interprétation n'est pertinente que « par son identité matérielle, phonique, sans aucun élément supplémentaire de contenu » (Gary-Prieur & Noailly, 2019 : 47), autrement dit, l'interprétation discursive reste restrictive.

2.3.2.2 Type métaphorique

Contrairement aux emplois prototypiques et dénominatifs des Npr dont l'objet d'interprétation est l'image initiale du référent en tant que tel, certains SN présentent un objet différent du référent initial du Npr employé : il s'agit alors d'une référence métaphorique et métonymique. Nous commençons par l'emploi métaphorique.

Les Npr peuvent être employés de façon métaphorique pour exprimer une relation d'analogie entre deux entités différentes. Pour cette raison, les interprétations métaphoriques exigent « une information minimale sur le référent » (Gary-Prieur & Noailly, 2019 : 51). Ces emplois sont souvent considérés comme des « pures créations discursives » parce qu'ils créent une relation analogique à partir d'un trait du référent initial. Par exemple, une expression comme « Pierre est un Sherlock Holmes » est comprise sans difficulté parce que le locuteur et son interlocuteur établissent un lien entre Pierre et Sherlock Holmes basé sur une caractéristique commune au référent et au référé, à savoir la qualité de détective telle que le sens de l'observation ou les capacités à raisonner, *etc.*

Jonasson (1994 : 215) insiste sur le fait que cet usage ne semble avoir de « restrictions ni en ce qui concerne le type de déterminant qui introduit le Npr, ni en ce qui concerne son nombre ». Les Npr métaphoriques peuvent être utilisés avec ou sans déterminant¹³ au singulier ou au pluriel comme nous venons de le montrer avec le cas de « Sherlock Holmes » *supra* ; une construction telle que « Ils sont tous des Sherlock Holmes » est aussi acceptable

¹³ Lorsque le Npr métaphorique est en position d'attribut ou d'apposition, il est susceptible d'être employé sans déterminant.

ex. Mel Brooks, *Don Quichotte du rire*,... (exemple cité par Jonasson, 1994 : 214).

lorsque nous voyons un groupe d'enfants qui essaient de trouver le coupable dans un dessin animé. Les Npr peuvent être accompagnés de divers types de compléments : les SN abritant ces Npr contiennent souvent des compléments modificateurs comme les adjectifs épithètes (par exemple, Pierre est un futur Sherlock Holmes) et les compléments adnominaux (par exemple, Pierre est le Sherlock Holmes de son village).

Selon la description de l'auteur, « au lieu de désigner un particulier porteur du Npr, le SN abritant le Npr signale maintenant un type ou un rôle » (*ibid* : 214), c'est-à-dire le référent de ce SN montre une certaine similarité ou joue le rôle de celui du Npr. L'emploi de ce type est donc descriptif : le référent est défini par des traits caractéristiques ou par le rôle du Npr en question supposé connu. C'est par l'intermédiaire de cette indication que le référent discursif de l'énoncé est identifié. Pour désigner cette fonction à la fois descriptive et identifiante, Lyons (1978)¹⁴ utilise l'expression d'« arrière-plan descriptif d'un nom propre » : cette fonction a pour but d'exprimer les faits essentiels et déjà établis sur le Npr et son porteur. Regardons les phrases suivantes :

(8) a. BTS, **les Beatles** de la génération YouTube (Le Figaro, 18/10/2018)

b. A Mâcon, il crée la « Grizi Académie », une école de football pour que les rêves des jeunes enfants soient exaucés et que **les futurs Griezmann** ne s'échappent plus. (L'OBS, 11/11/2020)

c. Lui, c'est **un Joker**.

Dans ces exemples, ce dont il faut tenir compte dans les emplois métaphoriques, c'est la valeur qualificative des référents discursif et initial. Pour interpréter (8a), nous utilisons l'arrière-plan descriptif du Npr : pour la jeune génération dite « la génération YouTube », le groupe de musique « BTS » est comme « les Beatles » faisant référence à un groupe de chanteurs mondialement connu pour la génération précédente. « Les jeunes enfants » qui vont à « Grizi Académie » dans (8b) sont décrits comme « les futurs Griezmann », le fameux joueur de football français. Ensuite, lorsque nous disons du référent de (8c) qu'il est « un Joker », nous estimons que celui-ci possède un caractère méchant ou joue le rôle du vilain comme le personnage « Joker » dans le film « Batman ».

Comme nous venons de le voir *via* l'exemple de « Sherlock Holmes » et celui de (8), le Npr peut être la tête de SN de plusieurs types : un Npr employé seul, un SN défini et un SN

¹⁴ Cité par Jonasson (1994 : 219).

indéfini¹⁵. Toutes ces constructions reposent sur les caractéristiques ou le rôle symbolique qu'indique le Npr initial, et elles établissent un lien entre les deux référents qui ne sont pas liés directement l'un à l'autre. L'interprétation métaphorique dépend ainsi à la fois des informations linguistiques données par le contexte et des connaissances partagées par les interlocuteurs.

2.3.2.3 Type métonymique

Les emplois métonymique et métaphorique du Npr ont pour point commun de renvoyer à un autre référent que le référent initial. Cependant, ils se distinguent par le fait que le Npr métonymique réfère à « un objet qui est en relation de contiguïté avec lui » (Gary-Prieur & Noailly, 2019 : 48) à la différence de l'emploi métaphorique fondé sur une ressemblance du contenu.

Gary-Prieur & Noailly (2019 : 48) distinguent deux cas de Npr métonymique : les Npr de personnes et les Npr de lieu. Les premiers s'emploient pour désigner un ou des objets appartenant à leurs référents initiaux¹⁶ alors que les derniers réfèrent aux produits fabriqués dans des lieux indiqués par les Npr comme *Champagne* ou *Camembert*¹⁷. Voici quelques exemples :

- (9) a. Les musiciens jouent **du Charles Aznavour** dans le métro.
- b. J'ai offert **le dernier Milan Kundera** à ma sœur. **C'**est mon roman préféré.
- c. Après qu'un demi-siècle et l'Atlantique les eurent séparés, Egon Schiele et Jean-Michel Basquiat se rencontrent à la Fondation Louis Vuitton, à Paris, jusqu'au mois

¹⁵ Un SN de la forme « un Npr » est parfois ambigu. Il peut être catégorisé comme métaphorique ou exemplaire. Ce dernier ne décrit pas un référent autre que son référent original contrairement à l'emploi métaphorique utilisé pour désigner un autre référent que le porteur original. Le type « exemplaire » se focalise sur le Npr historique ou notoire d'un référent bien connu (Jonasson, 1994 : 229-230). Mais il est parfois difficile de décider si « un Npr » renvoie au référent initial du Npr ou à un autre référent lorsqu'il est employé dans le texte. Jonasson (1994 :232) constate également que l'emploi exemplaire est « un premier pas vers la métaphorisation ».

Dans cette étude, nous nous contenterons de mentionner ce point délicat sans distinguer ces deux emplois sans rapport direct avec notre travail.

¹⁶ Il s'agit souvent des objets culturels signés par leurs auteurs.

¹⁷ Certain Npr de lieu se lexicalisent comme Nco, par exemple, certains produits nommés par leurs régions d'origine comme le camembert, le cognac, etc.

de janvier 2019. Attention, attendez-vous à **beaucoup de Basquiat** pour **peu de Schiele**. — (*MU in the city*, 04/12/2018)¹⁸

d. J'ai acheté **un Camembert** mais je n'ai pas acheté **le Roquefort**.

e. Si tu veux, il y a **du Cognac**.

Dans les deux cas, les Npr peuvent s'employer au singulier ou au pluriel. Ils sont considérés « comme massif si le déterminant est *du*, comme comptable dans tous les autres cas ». Ainsi, il ne semble pas exister de restrictions sur le nombre du déterminant. Cependant, une contrainte pèse sur tous les emplois de ce type de Npr : le déterminant doit avoir « une forme de masculin, quel que soit le genre du référent en question et le sexe du référent initial » (Gary-Prieur et Noailly, 2019 : 48).

Dans l'exemple (9a-c), le Npr ne désignant pas directement son porteur peut représenter un objet associé : *du Charles Aznavour*, c'est « des compositions musicales de Charles Aznavour », *le dernier Milan Kundera*, « le dernier livre de Milan Kundera », et *beaucoup de Basquiat* et *peu de Schiele* représentent respectivement « beaucoup de tableaux de Basquiat » et « peu de tableaux de Schiele ». Il en va de même pour (9d) et (9e). « Un Camembert », « le Roquefort » et « le Cognac » ne renvoient pas aux régions mais aux produits nommés d'après leur région. La relation entre le Npr du porteur et l'entité associée « préexiste à l'énonciation » (Riegel *et al.*, 2018 : 954). Pour interpréter cette sorte d'emploi, on doit disposer de certaines informations sur le référent initial du Npr. Mais, même sans idée sur la relation préexistante à l'énonciation, grâce à la prédication ou au contexte, on infère la classe du référent.

2.3.2.4 Type dit « de fractionnement »

Les Npr modifiés peuvent renvoyer à des référents « particuliers de particuliers ». En d'autres termes, ils peuvent désigner des « entités qui sont elles-mêmes des occurrences de particuliers ». Il s'agit d'emplois de fractionnement¹⁹ (Kleiber, 1996 : 580). Pour les décrire, Gary-Prieur (1991 : 50) insiste sur le fait que ces Npr, surtout lorsqu'ils sont employés avec

¹⁸ <https://www.mu-inthecity.com/un-peu-de-schiele-beaucoup-de-basquiat>

¹⁹ Il y a plusieurs étiquettes différentes pour cet emploi telles que « le type manifestation » (Jonasson, 1994) et « l'interprétation-image » (Gary-Prieur, 2016). Pourtant, nous utilisons le terme « fractionnement » pour une compréhension immédiate.

l'article défini, sont considérés comme référant à des « aspects » ou « manifestations » de l'individu initial et doivent être compris comme des référents existant dans un certain univers au lieu d'être entendu comme des fractions du référent. En d'autres termes, cet emploi ne concerne pas des morceaux du référent initial, mais bien ses images, nos façons de voir ce référent. Cependant, dans cette étude, nous ne traiterons pas de cette différence notionnelle : nous nous contenterons de souligner que, dans l'usage de ce type de Npr, la multiplicité des éléments ou des images, quel que soit le terme, composant le référent initial doit être présumée et que ces Npr ont naturellement la forme « le Npr Expansion » pour préciser l'élément ou l'image en question. Les SN comportent un adjectif (10a) et (10c) ou un complément adnominal introduit par *de* (10b) afin de construire le référent discursif :

(10) a. La force de **Hulk calme** est inférieure de celle de **Hulk enragé**.

b. **Bruce Wayne de la vie quotidienne** est un vrai gentleman.

c. **Le jeune Camus**, encore en Algérie, s'enthousiasme, en 1938, pour le premier roman de Sartre, *La Nausée*.

Dans ces énoncés, nous parlons des instances de particuliers, « Hulk », « Bruce Wayne » et « Camus ». Ces Npr indiquent respectivement une fraction spécifique des référents. Autrement dit, ils soulignent une certaine facette du référent. Par conséquent, dans (10a), les adjectifs « calme » et « enragé » montrent les différents états du personnage imaginaire, « Hulk ». Il en va de même dans (10b). L'information fournie par le complément adnominal présente un aspect spécifique du personnage. L'adjectif « jeune » de (10c) participe d'une saisie référentielle du référent « Camus » à une période de sa vie : la jeunesse. Comme le confirme la *GMF*, ce type de SN « opère une scission ou une division de l'individu porteur du nom propre » (Riegel *et al.*, 2018 : 338-339).

(11) **un Stade de France stupéfait** au coup de sifflet final

L'exemple (11) montre un double effet du Npr modifié. Le SN « un Stade de France stupéfait » indique un moment où le sentiment règne sur l'espace entier. Cette expression peut être catégorisée comme l'emploi de fractionnement parce qu'elle indique un moment d'existence de cet être, « Stade de France ». En même temps, cet exemple est métonymique : un lieu qui est un référent inanimé, lui-même, ne peut pas avoir le sentiment. Par conséquent, ce SN renvoie au public se trouvant dans le « Stade de France » à un moment bien précis.

2.4 Les emplois des noms propres en coréen

Nous venons de présenter les différents *Npr* et leurs divers emplois classés selon leurs traits syntaxique et sémantique. Dans les pages suivantes, en restant dans le cadre de l'universalité du phénomène linguistique, nous introduirons la manière dont les différences structurales de la langue coréenne par rapport au français sont manifestées dans les utilisations des *Npr*.

2.4.1 La définition des *noms propres* en linguistique coréenne : l'emploi prototypique

Traditionnellement, dans la grammaire coréenne, les *Npr* sont considérés comme des expressions désignant des êtres uniques qui ne peuvent pas s'employer au pluriel (Nam K. S. et *al.*, 1985 ; Lee K. K., 2002 ; Lee J. T., 2011 ; Yoo H. K. et *al.*, 2015). Un certain nombre de linguistes coréens soulignent que la principale distinction entre les *Npr* et les *Nco* est la possibilité, pour ces derniers, d'avoir des référents pluriels (Lee J. T., 2011 : 272-273) :

« Les noms communs sont les noms généraux renvoyant aux êtres existants en quantité et les *Npr* renvoient à un seul être. » (Lee K. K., 2002 : 127)

« Les noms propres sont des noms attribués aux êtres particuliers et chaque être particulier est unique. » (Nam K. S. et *al.*, 1985 : 46)²⁰

Cependant, à notre avis, il s'agit d'une perspective qui ne dépasse pas la définition traditionnelle : le nombre de référents n'est pas un critère suffisant pour différencier les *Npr* et les *Nco*. Voici un fameux contre-exemple : le soleil et la lune, qui renvoient à des êtres uniques, sont toutefois des *Nco*. Contrairement à ces derniers, les *Npr* peuvent avoir plusieurs porteurs comme le nom de famille « Kim » qui est très répandu en Corée. Ainsi, il est difficile d'affirmer que les *Npr* renvoient seulement à des êtres uniques. Leur définition proposée dans la linguistique coréenne qui met l'accent sur la condition d'unicité du référent n'est pas suffisante pour décrire le phénomène et donc sera exclue de notre étude.

Dans cette optique, nous soulignerons, dans les pages suivantes, que les *Npr* peuvent être classés en deux sous-catégories comme nous l'avons démontré dans la partie précédente : les *Npr* non modifiés et les *Npr* modifiés comprenant un déterminant ou des modificateurs. Etant donné l'absence d'article, il ne paraît pas adéquat à première vue de

²⁰ Les deux citations ont été traduites directement par nos soins.

transposer ce classement en coréen. En effet, cette langue possède d'autres dispositifs grammaticaux qui jouent un rôle proche de celui de l'article français²¹.

Nous pouvons facilement déduire ces dispositifs par le biais de « trois interdictions dans les emplois des Npr » (Nam K. S. et *al.*, 1985 : 47) théoriquement proposées par la grammaire coréenne :

- Les Npr ne s'emploient pas avec le suffixe indiquant la pluralité « -deul » et la particule auxiliaire²² exprimant l'individualité « -mada » :

- Le suffixe « -deul » s'ajoute à la fin d'un mot pour marquer la pluralité du référent²³. L'exemple suivant montre le rôle du suffixe « -deul » lorsqu'il est employé avec un Nco :

ex) 거리에 사람들이 있다.
geoli+e salamduel+i iss+da.
rue+cas-adv. personne+suf.pl+cas-sujet il y a+terminaison.
Il y a **des gens** dans la rue.

Dans cet exemple, le suffixe « -deul » employé avec le Nco « saram (personne) » indique qu'il y a plusieurs personnes dans la rue.

- Ensuite, la particule auxiliaire « -mada » est sémantiquement équivalente à « chaque » en français et montre l'individualité du référent avec lequel elle s'emploie. Regardons l'exemple suivant :

ex) 거리마다 사람들이 있다.
geoli+mada salamduel+i iss+da.
rue+aux.ind personne+suf.pl+cas-sujet il y a+terminaison.

²¹ Voir la section 2.2 du chapitre 3.

²² La particule auxiliaire s'agglutine à la fin du mot pour lui conférer le sens prépositionnel ou adverbial (voir la section 3.2.1 du chapitre 3).

²³ En règle générale, la fréquence de l'utilisation de la marque de pluralité (par exemple « -deul ») en coréen est inférieure à celles de marque « -s » en français et en anglais. L'emploi de la marque de pluralité n'est pas obligatoire et il n'est parfois pas naturel d'ajouter « -deul » même à la fin du Nco. Par exemple, certains noms inanimés et des noms concrets tels que « pomme » et « café », les noms abstraits « amour » ne s'emploient pas avec « -deul ». Mais il est pertinent de dire que, de manière générale, la marque du pluriel peut s'employer avec le Nco.

Il y a des gens dans **chaque rue**.

- Les Npr ne s'emploient pas avec *gwanhyungsa* :
 - *Gwanhyungsa* est une sous-catégorie d'épithète²⁴ qui se place devant les noms pour les qualifier. Il s'agit d'une catégorie grammaticale invariable, dépendante et restreinte aux noms. Elle se divise en trois sortes : *seongsang gwanhyungsa* décrivant la nature ou l'état des choses comme *sae (nouveau)*, *heon (ancien)*, *sun (pur)*, *ju (principal)*, etc., *su-gwanhyungsa* (adjectif numéral) et *jisi gwanhyungsa* (déterminant démonstratif)²⁵. D'après la définition traditionnelle des *Npr* de la grammaire coréenne (Nam Kisim et al., 1985 : 169-173), il est difficile de construire un SN de forme « *gwanhyungsa* + *Npr* ».
- Les Npr ne s'emploient pas avec les mots qui concernent le nombre :
 - A part *su-gwanhyungsa* équivalent à l'adjectif numéral en français, *susa* est un nom indiquant le nombre²⁶.

Les dispositifs grammaticaux mentionnés ci-dessus expriment l'individualité, la pluralité ou même la « définitude » du référent et jouent un rôle approximativement équivalent à celui des déterminants. Ces trois propositions concernent donc les emplois prototypiques des *Npr* non-modifiés. Pourtant, l'objectif de cette partie est d'analyser les emplois des *Npr* coréens et leurs effets linguistiques diversifiés avec un spectre plus large par rapport à la grammaire traditionnelle coréenne afin de fonder une base de comparaison entre le français et le coréen, et de fournir à la linguistique coréenne une possibilité de redéfinir la notion de *Npr*. Dans cette optique, nous élargirons notre point de vue pour inclure également les *Npr* modifiés dans le domaine d'analyse des *Npr*.

Nous estimons que les éléments mentionnés par ces propositions, le suffixe « -deul », la

²⁴ Comme mentionné dans la section 2.2 du chapitre 3 (la note 18), *Gwanhyungsa* et *hyungyongsa* sont tous deux des expressions épithètes. Cependant, la première est invariable et se situe toujours devant les noms alors que la deuxième est un mot variable qui peut se trouver non seulement devant un nom mais aussi en position du prédicat en exprimant l'état du sujet.

²⁵ La troisième sorte de *gwanhyungsa* qui a un fort caractère démonstratif sera principalement traitée dans les pages ultérieures avec les autres cas de SN démonstratif (voir la section 4.4.2 de ce chapitre).

²⁶ Nous ne les présenterons pas ici parce que si nous arrivons à montrer que les *Npr* peuvent être employés avec *su-gwanhyungsa*, cela suffit à réfuter la troisième interdiction, « les *Npr* ne s'emploient pas avec les mots qui concernent le nombre ».

particule auxiliaire « -mada », *gwanhyungsa*, et, finalement, les mots qui concernent le nombre, peuvent participer à la construction des SN dont la tête est le Npr, à savoir les Npr modifiés. Ainsi, nous appliquerons au coréen le classement des emplois des Npr non-prototypiques tel que présenté dans la section 2.3.2 *supra*. Sur la base de ce classement, notre objectif dans la partie suivante consistera à présenter chaque emploi des Npr utilisés avec les éléments grammaticaux considérés comme inacceptables en coréen dans le cadre des emplois prototypiques des Npr.

2.4.2 Les emplois non-prototypiques

2.4.2.1 Type dénominatif et éléments grammaticaux

Nous avons constaté plus haut que les Npr dénominatifs définissent une entité par son appartenance à l'ensemble de référents portant le même nom ou bien à introduire une entité pour l'interlocuteur. Par ailleurs, en français, ces Npr n'ont pas des restrictions formelles sauf la forme du Npr défini singulier employé seul (« le Npr »). Dans cette partie, notre objectif est de voir si les Npr coréens de ce type peuvent s'employer avec des dispositifs grammaticaux mentionnés *supra*, le suffixe « -deul » et la particule auxiliaire « -mada », *gwanhyungsa* y compris les mots concernant le nombre.

- Le suffixe « -deul » et la particule auxiliaire « -mada »

Premièrement, le suffixe « -deul » peut être employé avec les Npr afin d'exprimer la pluralité lors d'un emploi dénominatif :

(12) 김씨들이	모여	사는	마을
Kimssi <u>deul</u> +i	moyeo	saneun	maeul
Kim+affixe ²⁷ + <u>suf.pl</u> +cas-sujet	se rassembler	habiter	village
Le village où habitent les Kim . (village clanique des Kim)			

Nous avons déjà emprunté la notion de « prédicat de dénomination » de Kleiber (1981) pour définir l'emploi dénominatif des Npr (voir la section 2.3.2.1). Par ailleurs, nous avons remarqué que la formule « être appelé Npr » peut se réaliser non seulement au singulier et au pluriel mais aussi sous la forme d'un SN indéfini ou défini (sauf le défini

²⁷ « -ssi » est un affixe uniquement employé avec le patronyme et indique ce qui précède est un patronyme ou une famille de tel patronyme.

singulier).

Dans l'exemple (12) qui concerne l'emploi du patronyme²⁸, « Kim », employé au pluriel à l'aide de « -deul », indique que plusieurs personnes portent le même nom, et cela est parfaitement acceptable.

Ensuite, si le référent d'un Npr est inéluctablement pluriel, par exemple, lorsqu'il s'agit d'un produit commercial²⁹, d'un rôle d'une pièce de théâtre joués par plusieurs personnes ou des référents possédant un caractère périodiquement répétitif, les Npr peuvent être employés avec la particule auxiliaire « -mada » de manière dénominative comme dans les exemples (13) et (14) :

(13) 아이폰마다 고유번호가 있다.

Iphone+**mada** goyubeonho+ga iss+da.

Iphone+**chaque** numéro de série+cas-suj. il y a+term.

Chaque Iphone a un numéro de série.

(14) 나는 크리스마스마다 여행을 간다.

²⁸ Ariel (1990 : 44) distingue, parmi les Npr de nom de personne, les différents degrés de l'accessibilité entre le prénom, le patronyme, et le nom complet : les prénoms sont les plus équivoques, les patronymes sont des intermédiaires, et enfin les noms complets ont l'accessibilité la plus basse, c'est-à-dire ces derniers portent plus d'informations lexicales parmi ces trois. Nous sommes d'accord sur le principe de cette distinction, mais comme notre objectif de ce chapitre est de fournir une description globale sur les Npr présentés dans les deux langues, nous ne traitons pas ces différences entre les noms de personne soulignées par l'auteur.

²⁹ Pour les noms commerciaux (nom de marque et nom de produit), quelques points nécessitent une discussion parce que les frontières qui distinguent certains noms enregistrés dans les dictionnaires et considérés comme des noms communs et ceux qui restent comme Npr ne sont pas toujours nettes. Selon la description sémantique de Fèvre Pernet & Roché (2005 : 13), les noms de produit, surtout fabriqués en grande quantité, ont un fort caractère du Nco et sont donc sémantiquement analysables alors que les noms de marque sont plutôt des « désignateurs rigides ». Cependant, certains noms de marque remplacent tous les produits de la même classe. Dans ce cas-là, la représentativité prend une place importante parce que si le produit n'a pas un certain niveau de représentativité parmi les autres de la même catégorie, on aura des difficultés à inférer la partie réduite, à savoir la classe du porteur. Par exemple, dans « Elle a acheté un Post-it », le SN « un Post-it » ne pose aucun problème et renvoie directement à la note repositionnable. Ce nom de marque ayant une haute représentativité parmi ses semblables, est quasiment devenu un nom commun pour ce type de produits quels que soient leurs marques. Par cela, nous pouvons dire, bien évidemment avec prudence, qu'il se peut que l'appartenance ou non au lexique des noms commerciaux est une question de degré plutôt qu'une question de nature. Néanmoins, même si nous reconnaissons la plausibilité de ces arguments, dans notre recherche, nous décidons de ne pas faire la différence entre ces différents degrés et considérer tous les noms commerciaux comme des Npr.

na+neun Christmas+**mada** yeohaeng+eul ga+n+da.
 Je+cas-suj. Noël+**chaque** voyage+cas-obj. partir+présent+term.
 Je pars en voyage **chaque Noël**.

● *gwanhyungsa*

Les Npr de type dénominatif sont susceptibles d'indiquer, pour emprunter le terme de Gary-Prieur & Noailly (2019 : 43), « la pluralité interne »³⁰ et, dans tels cas, ils désignent des individus qui font partie d'un ensemble dénoté par un nom collectif. Ils peuvent donc s'employer avec *su-gwanhyungsa* (numéral) et *jisi gwanhyungsa* (déterminant démonstratif).

D'abord, *su-gwanhyungsa* équivalant aux mots numéraux (*un, deux, trois...*) ne peut pas être utilisé directement avec les Npr. En règle générale, pour participer à la construction d'un SN, *su-gwanhyungsa* s'accompagne d'un nom dépendant (Ndép)³¹ qui a une fonction de classificateur et qui indique la classe et la quantité du référent d'un Npr comme le montre le SN de (15b) :

(15) a. *두 포르세

³⁰ Il s'agit d'une appréhension collective de la référence (voir la note 12 *supra*).

³¹ Les noms dépendants (abrégié Ndép) sont non-autonomes et ne peuvent pas être utilisés seuls. Ils s'emploient souvent avec diverses particules de cas et sont classés en plusieurs sous-catégories selon leur fonction : les Ndép universels, les Ndép subjectifs, les Ndép adverbiaux, les Ndép prédicatifs et les Ndép classificateurs.

- Les Ndép universels : certains Ndép renvoient à une personne, une chose, un lieu ou une affaire.
 ex) meogneun *geos* (chose à manger) / meogeul *gos* (lieu pour manger).
- Les Ndép subjectifs : les Ndép qui se combinent avec la particule de cas-sujet et jouent le rôle du sujet dans la phrase. ex) Seouleul tteonan *jiga* (depuis que je suis parti de Seoul).
- Les Ndép adverbiaux : ceux qui jouent le rôle d'adverbe lors de l'utilisation avec la particule de cas-adverbe ex) sinbal sineun *chaelo* deuleoojima (N'entre pas en mettant tes chaussures).
- Les Ndép prédicatifs : ils s'emploient avec la particule de cas-prédicat, *-ida*, et jouent le rôle du prédicat dans la phrase. ex) ineun somunil *ppunida* (ce n'est qu'une rumeur).
- Les Ndép classificateurs : certains mots montrent le fait qu'il s'agit de la quantité du référent concerné et aussi le type de ce référent. ex) gangaji du *mari* (deux chiots).

Ces derniers fonctionnent comme classificateurs qui fournissent l'information sur le type de référent des noms qu'ils accompagnent : le choix du Ndép diffère selon le nom accompagné. Par exemple, pour les noms désignant les transports, on utilise toujours le Ndép « dae » et les noms renvoyant aux animaux s'accompagnent du Ndép « mari ». Les Ndép classificateurs participent à la construction d'un SN en exprimant l'intention locutionnelle de préciser la quantité du référent et définissant le type de référent, à savoir, la catégorie d'appartenance du référent.

du Porsche
deux-suGHS Porsche

b. 두 대의 포르세
du dae+ui Porsche
deux-suGHS classificateur-voiture+cas-posséssif Porsche
Deux Porsche

A part les mots numériques directs qui désignent les chiffres précis, les expressions exprimant la pluralité telles que « tout », « plusieurs », *etc.* appartiennent aussi à la catégorie de *su-gwanhyungsa*. Dans ce cas-là, elles peuvent être ajoutées directement, sans Ndép, aux Npr de fonction dénomminative :

(16) 모든 김씨가 다 모였다.
modeun Kimssi+ga da moyeo+ss+da.
tout-suGHS Kim+Ndép+cas-suj. tout-adv. se rassembler+passé+term.
Tous les Kim se sont rassemblés.

Ensuite, les *jisi gwanhyungsa* sont également susceptibles de s'employer avec les Npr de ce type. La plupart des *jisi gwanhyungsa* se caractérisent comme déictiques³² et quelques-uns expriment une indétermination : *certain, quelque, autre, etc.* Ces dernières unités³³ peuvent être utilisées avec les Npr pour réaliser leur fonction dénomminative :

(17) 어떤 김서영이 널 찾았어.
eotteon KimSeoYeoung+i neol chaja+ss+eo.
Certain KimSeoYeoung+cas-suj. te chercher+passé+term.
Une certaine SeoYeoung Kim t'a cherché.

2.4.2.2 Type métaphorique et éléments grammaticaux

A partir d'une relation analogique entre deux entités, les Npr peuvent participer à la construction métaphorique. Dans cet emploi, le Npr désigne un objet différent de l'image initiale du référent. En corollaire, ce type de Npr se montre sélectif sur les déterminants et les

³² Ils seront mentionnés dans la partie dédiée aux SN démonstratifs (cf. la section 4.3).

³³ Ce sont des mots considérés comme « adjectifs indéfinis » en français, donc, ils ne sont pas les démonstratifs. Pourtant en coréen, la plupart des adjectifs indéfinis qui assurent la pluralité sont catégorisés comme *su-gwanhyungsa* et ces mots en question sont classés comme *jisi gwanhyungsa*.

compléments qui faciliteraient sa fonction descriptive. Ici, nous examinerons la condition d'emploi des Npr métaphoriques en coréen.

- Le suffixe « -deul » et la particule auxiliaire « -mada »

Les Npr combinés avec « -deul » peuvent effectuer une interprétation métaphorique. Regardons l'exemple suivant :

(18) 거리에 **헐크들이** 있다.
geoli+e Hulkdeul+i iss+da.
rue+cas-adverbe Hulk+suf.pl+cas-sujet il y a+terminaison.
Il y a **des Hulk** dans la rue.

Cette phrase s'interprète de diverses manières selon la situation d'énonciation. Pourtant, pour une interprétation métaphorique, supposons une situation telle qu'il y a des gens bien musclés démontrant leur force physique comme le personnage Hulk. Dans ce cas-là, le SN de forme « Npr+deul » ne pose aucun problème.

Les SN de ce type peuvent avoir la forme « épithète+Npr+deul » comme illustré dans l'exemple (19) :

(19) 미래의 **그리즈만들이** 축구를 한다.
milae+ui Griezmanndeul+i chuggu+leul ha+n+da.
Futur+cas-posséssif Griezmann+suf.pl+cas-suj. foot+cas-obj faire+présent+term.
Les futurs Griezmann jouent au foot.

Le Npr « Griezmann » qui normalement renvoie directement à une personnalité unique est ici employé avec une épithète et le suffixe de pluriel. Dans cette construction, il acquiert une valeur métaphorique³⁴.

Étant donné que les Npr de type métaphorique sont susceptibles de s'employer au pluriel, corollairement, ils peuvent exprimer l'individualité lorsqu'ils sont employés avec la particule auxiliaire « -mada ». Bien évidemment, cela exige une condition préalable sur la pluralité des

³⁴ Nam Kisim et al. (1985 : 47) insistent sur le fait que, lorsqu'un Npr est employé de cette façon, il devient Nco. Pourtant, de notre point de vue, les Npr gardent toujours leur propre valeur en tant que tels même dans les cas où ils sont utilisés métaphoriquement. Autrement dit, pour réaliser une métaphore, une connaissance précise du contenu du Npr, une connaissance « des propriétés singulières de son référent initial » est nécessaire (Gary-Prieur & Noailly, 2019 : 51). Nous estimons que les auteurs coréens ont proposé un tel point de vue à partir d'une définition restreinte sur les Npr qui exclut les Npr non-prototypiques.

Npr. Reprenons l'exemple (18) *supra.*, et ajoutons une précision comme suit :

(18') 거리에 헬크들이 있다.

geoli+e Hulkduel+i iss+da.

rue+cas-adverbe Hulk+suf.pl+cas-sujet il y a+terminaison.

헬크마다 조금씩 다르게 생겼다.

Hulkmada jogeumssig daleuge saenggyeo+ss+da.

Hulk+chaque un peu différemment apparaître+passé+terminaison.

Il y a des Hulk dans la rue.

Chaque Hulk est un peu différent.

Avec une présupposition contextuelle soit explicite soit implicite, la particule auxiliaire peut s'employer avec un Npr de type métaphorique.

- *gwanhyungsa*

Les Npr de type métaphorique peuvent s'employer avec *su-gwanhyungsa*. Contrairement à l'exemple (15) où le Npr fonctionne dénominalement et donc exige à la fois un Ndép et *su-gwanhyungsa* pour indiquer la quantité du référent, le Npr de ce type est utilisé directement avec *su-gwanhyungsa* comme illustre l'exemple (20) :

(20) 두 헬크의 대결

du Hulk+ui daegyeol

deux-suGHS Hulk+cas-possessif confrontation

Un confrontation entre deux Hulk

2.4.2.3 Type métonymique et éléments grammaticaux

Nous avons constaté que les Npr métonymiques s'emploient afin de désigner un autre référent que son référent initial comme le font les Npr métaphoriques. Cependant, le type métonymique s'appuie sur la relation de contiguïté entre des entités. Nous examinerons si la combinaison des Npr avec les dispositifs grammaticaux en coréen mentionnés *supra* serait susceptible de réaliser l'effet métonymique.

- Le suffixe « -deul » et la particule auxiliaire « -mada »

En coréen, les Npr de type métonymique ne se servent ni du suffixe « -deul » ni de la particule auxiliaire « -mada ».

- (21) a. * 나는 베토벤들을 자주 들어
 * na+neun Beethovendeul+eul jaju deul+eo.
 * Je+cas-sujet Beethoven+suf.pl+cas-objet souvent écouter+terminaison.

- b. 나는 베토벤을 자주 들어.
 na+neun Beethoven+eul jaju deul+eo.
 Je+cas-sujet Beethoven+cas-objet souvent écouter+terminaison.
 J'écoute souvent **du Beethoven**.

Lorsque nous voulons désigner plusieurs entités du référent d'un Npr métonymique, en coréen, l'utilisation simple d'un Npr est nécessaire comme dans (21b).

En outre, la particule auxiliaire « -mada » ne s'emploie pas directement avec le Npr métonymique. Afin d'exprimer le caractère singulier d'un référent, « -mada » s'ajoute au Nco qui suit le Npr et précise la classe du référent comme le montre (22a-b) :

- (22) a. * 베토벤마다 특색이 있다.
 * Beethovenmada teuksaeg+i iss+da.
 * Beethoven+chaque particularité+cas-objet avoir+terminaison.

- b. 베토벤 곡마다 특색이 있다.
 Beethoven gok+mada teuksaeg+i iss+da.
 Beethoven œuvre+chaque particularité+cas-objet avoir+terminaison
Chaque œuvre de Beethoven a une particularité.

● *gwanhyungsa*

En règle générale, en coréen, il est difficile pour les Npr de type métonymique d'accompagner *gwanhyungsa*. La langue coréenne a tendance à utiliser le SN de forme « Npr+*gwanhyungsa*+Nco » et « Npr+*su-gwanhyungsa* (chiffre)+Ndép » comme dans les exemples qui suivent :

- (23) a. * 새 쿤데라를 샀다.
 * sae Kunderareul sa+ss+da.
 * Nouveau-ssGHS Kundera+cas-obj. acheter+passé+terminaison.

- b. 쿤데라 새 책을 샀다.
 Kundera sae chaeg+eul sa+ss+da.
 Kundera nouveau-GHSn livre(NomC)+cas-obj. acheter+passé+terminaison.

(J'ai) acheté **le nouveau Kundera**.

(24) a. *두 쿤데라를 샀다.

*du **Kundera**+reul sa+ss+da.

*Deux-suGHS **Kundera**+cas-obj. acheter+passé+terminaison.

b. 쿤데라 두 권을 샀다.

Kundera du gwon+eul sa+ss+da.

Kundera deux-suGHS Ndép(livre)+cas-obj. acheter+passé+terminaison.

(J'ai) acheté **deux Kundera**.

Comme nous le voyons dans (23a) et (24a), les Npr de type métonymique ne s'emploient pas directement avec *gwanhyungsa*. Un nom soit Ndép soit Nco qui peut préciser la classe du référent désigné par le Npr en question est nécessaire afin de réaliser le sens métonymique comme illustré dans (23b) et (24b).

2.4.2.4 Type dit de fractionnement et éléments grammaticaux

Les Npr de fractionnement concernent la désignation d'une partie, d'une facette de particuliers. Une extension formelle de Npr est exigée dans les emplois de ce type afin de préciser le référent. Dans cette partie, nous montrerons la corrélation entre cette sorte d'usage et les dispositifs grammaticaux mentionnés en coréen.

- Le suffixe « -deul » et la particule auxiliaire « -mada »

Les Npr de fractionnement ne sont pas susceptibles d'accompagner le suffixe de pluralité « -deul » : cette sorte de Npr présuppose une scission du référent. En d'autres termes, le référent finalement désigné par le SN est une partie spécifique du référent initial qui est un particulier unique. Par conséquent les Npr de ce type ne peuvent pas être mis au pluriel comme le présente l'exemple (25) :

(25) 바젤의 라인강과 암스테르담의 라인강.

Basel+ui **Rheingang**+gwa Amsterdam+ui **Rheingang**.

Bâle+cas-poss. **Rhin(rivière)**+et Amsterdam+cas-poss. **Rhin(rivière)**³⁵.

³⁵ En coréen, certains Npr, surtout les patronymes et les toponymes de nature, ont souvent une forme de mot composé : ces Npr sont accompagnés d'un autre mot, soit un affixe soit un Nco qui définit le type de référent. Par

*라인강들의 풍경이 다 다르다.

Rheingang+deul+ui punngyeong+i da daleu+da

Rhin-rivière+suf.pl+cas-poss. paysage+cas-suj. tout différent+term.

Le Rhin de Bâle et le Rhin d'Amsterdam. *Les paysages **des Rhin** sont différents³⁶.

Dans la même perspective, comme le SN comprenant le Npr fait partie d'un tout et comme son référent est un « particulier d'un particulier », il serait ironique de le mettre avec la particule auxiliaire « -mada » qui a le sens équivalent à « chaque » comme illustre l'exemple suivant :

(26) 바젤의 라인강과 암스테르담의 라인강.

Basel+ui **Rheingang+gwa** Amsterdam+ui **Rheingang.**

Bâle+cas-poss. **Rhin-rivière+et** Amsterdam+cas-poss. **Rhin-rivière.**

*라인강마다 풍경이 다 다르다.

Rheingang+mada punngyeong+i da daleu+da.

Rhin-rivière+chaque paysage+cas-suj. tout différent+terminaison.

Le Rhin de Bâle et le Rhin d'Amsterdam. *Chaque Rhin a un paysage différent³⁷.

Ainsi, les Npr de type de fractionnement ne peuvent être employés ni avec « -deul » ni avec « -mada ».

- *gwanhyungsa*

Voici quelques exemples de Npr de fractionnement utilisés avec *gwanhyungsa* (voir la section 2.4.1 *supra*). Ce dernier décrit la nature, le caractère ou le statut d'une facette particulière du Npr :

(27) 옛 신라의 수도

yet Silla+ui sudo

ancien-ssGHS Silla+cas-possessif capitale

conséquent, en coréen, la correcte appellation pour « Rhin » est « Rhin+rivière ».

³⁶ Il s'agit d'une traduction littérale : nous admettons que la bonne traduction de cette phrase serait « Tous ces Rhin offrent des paysages extrêmement différents ». Pourtant, nous avons choisi la façon mot à mot afin de montrer que la phrase en question est agrammaticale en coréen.

³⁷ *Idem.*

La capitale de (l'**ancien**) Silla³⁸

- (28) 새 대한민국을 위한 논의
Sae Daehanminguk+eul wihan nonui
nouveau-ssGHS Corée du Sud+cas-obj. pour débat
Le débat pour **la nouvelle Corée du Sud**

Dans (27) et (28), « yet Silla (l'**ancien Silla**) » et « sae Daehanminguk (la nouvelle Corée du Sud) » sont des SN de forme « *seongsang gwanhyungsa* + Npr ».

Cependant, *su-gwanhyungsa* (adjectif numéral) ne semble pas susceptible d'être employé avec Npr lors d'un usage par fractionnement tandis que *jisi gwanhyungsa* déictique peut sûrement s'antéposer aux Npr de ce type³⁹.

2.5 Bilan

Pour clore cette partie sur les Npr, nous résumons nos principales observations en quatre points :

- Les Npr eux-mêmes n'ont pas de sens, mais un lien s'établit entre eux et leurs porteurs. Sur la base de ce lien, ils construisent le sens en rapport avec le contexte et la situation d'énonciation.
- Les emplois des Npr se subdivisent en deux sous-catégories : une prototypique et une autre non-prototypique. La première concerne les Npr purs alors que la deuxième renvoie aux SN dont la tête est le Npr, c'est-à-dire les Npr modifiés.
- L'utilisation non-prototypique des Npr peut encore comprendre quatre sous-catégories, les emplois dénominatif, métaphorique, métonymique et de fractionnement, selon l'effet linguistique réalisé.
- Considérant le caractère universel du langage humain, ces analyses sont également applicables au coréen. Cependant, comme le coréen ne possède pas d'article, plusieurs autres éléments linguistiques participent à la construction des SN en fonction de l'usage et de l'objectif de l'énonciation.

³⁸ Silla est un des royaumes historiques de Corée (57 AEC - 668 EC).

³⁹ Nous aurons l'occasion de préciser les cas de *jisi gwanhyungsa* déictique lorsque nous parlerons des SN démonstratifs.

Bien que nous estimions que notre argumentation contestant le point de vue le plus communément admis dans la linguistique coréenne, qui exclut les emplois non-prototypiques, pourrait susciter des objections, nous avons choisi d'adopter une conception étendue des Npr après avoir discuté de leurs caractéristiques et de leur relation avec le référent. En effet, l'idée d'exclure du débat des Npr qui ne sont pas utilisés prototypiquement contourne la complexité du sujet. De plus, comme notre étude a pour objectif de comparer deux langues complètement différentes, nous avons estimé qu'il était nécessaire de présenter les critères uniformes afin d'aboutir à des analyses significatives.

Notre conclusion préliminaire par rapport aux Npr est qu'ils peuvent s'employer soit seuls soit avec d'autres éléments linguistiques en fonction du contexte ou de la situation de l'énonciation, ce qui engendre différents effets linguistiques. Dans cette perspective, notre enjeu principal est d'identifier les référents appropriés auxquels les Npr se réfèrent réellement dans le contexte ou dans la situation d'énonciation.

3 Les pronoms personnels

La définition étymologique du terme *pronom* est « à la place d'un nom ». Donc, le pronom est traditionnellement considéré comme un mot remplaçant un nom. Cependant, cette définition est remise en cause en linguistique en raison du fait que les propriétés sémantiques et les fonctionnements référentiels des pronoms sont, en effet, plus complexes et diversifiés selon l'observation de l'usage. Les pronoms ne remplacent pas simplement les noms : ils peuvent être également employés à la place du syntagme nominal, du syntagme prépositionnel, de l'adjectif, de la proposition, *etc.* comme le montrent les phrases (29a-d) :

- (29) a. **Les enfants** font du vélo. **Ils** roulent lentement.
b. Il a déjà parlé **de son projet** ? – Oui, il **en** a parlé
c. Sinon j'aurais été encore plus **triste** que je ne **le** suis déjà.
d. **Elle fait ses études à Strasbourg.** – Je **le** sais.

En français, plusieurs sous-classes constituent les expressions pronominales : les pronoms personnels, démonstratifs, possessifs, indéfinis, interrogatifs, relatifs et réfléchis.

Cependant, dans cette partie, seuls les pronoms personnels qui font partie des pronoms définis seront traités⁴⁰. Contrairement aux SN, ils n'ont pas de tête lexicale et ne comportent pas « au moins directement, d'indication sur la catégorie de leur référent » (Charolles, 2002 : 183). À cet égard, et comme nous l'avons remarqué au début de ce chapitre, il est pertinent de dire que les pronoms ont un mode de saisie différent de celui des noms propres. Nous soulignons donc qu'ils ont leur propre manière de combler le déficit descriptif et que leurs caractéristiques et emplois devront être examinés sous cet angle.

3.1 Les formes morphologiques variées du pronom personnel

Plusieurs éléments grammaticaux expriment la *personne*. En français, elle peut être exprimée par des pronoms personnels, des désinences verbales, des adjectifs et des pronoms possessifs, *etc.* Dans cette partie, nous nous concentrerons seulement sur les pronoms personnels.

Comme nous l'avons déjà mentionné plus haut, les pronoms sont des termes grammaticaux dont la forme diffère selon la fonction grammaticale. Ainsi, les pronoms personnels sont traditionnellement classés comme suit en fonction de leurs formes morphologiques :

Rang	Nombre	Personne	Formes conjointes			Formes disjointes
			Sujet	Complément direct	Complément indirect	
1	Sing.	1 ^e	Je	Me		Moi
2		2 ^e	Tu	Te		Toi
3		3 ^e	Il / Elle	Le / La	Lui	Y / En
			On	Se		Soi(-même)
4	Plur.	1 ^e	Nous	Nous		
5		2 ^e	Vous	Vous		
6		3 ^e	Ils / Elles	Les	Leur	Y / En
			Se			Eux / Elles (-mêmes)

Tableau 6 Catégories et variations morphologiques (GMF, 2018 : 368)

Ce tableau nous permet d'ébaucher une conception globale sur les caractéristiques des

⁴⁰ Les autres pronoms participant à la constitution des SN seront mentionnés partiellement et respectivement dans les parties SN définis, SN indéfinis et SN démonstratif.

pronoms personnels. Trois traits sont à remarquer. Premièrement, du point de vue syntaxique, nous distinguons des formes conjointes et d'autres disjointes (Riegel *et al.*, 2018 : 368-374). Les formes conjointes sont clitiques ; elles s'antéposent au verbe comme dans (30a) tandis que les formes disjointes sont plutôt syntaxiquement similaires aux SN qui sont séparés du verbe par une préposition ou une pause comme le montre (30b) :

(30) a. Elle **me** connaît.

b. Il se comprennent entre **eux**.

Ensuite, les formes conjointes se divisent en trois selon leur rôle dans la phrase : sujet, complément direct et complément indirect. Certaines formes de pronoms personnels se trouvent en place du sujet alors que certaines autres sont réservées à l'objet.

En dernier lieu, des formes différenciées existent pour les pronoms de troisième personne. Lorsqu'ils sont employés au singulier comme sujet ou complément direct, ils distinguent le genre. Lors d'un emploi en tant que complément indirect, ces pronoms différencient les humains et les non-humains.

3.2 La fonction référentielle des pronoms personnels

Charolles (2002 : 186) souligne que « les pronoms codent un mode de présentation et de donation de la référence qui leur est propre ». Dans cette section, nous introduirons les détails de ces codages variables selon le type de pronoms.

3.2.1 Les pronoms de première et de seconde personnes

Les pronoms de première et de seconde personnes (rang 1, 2, 4 et 5 qui sont *je*, *tu*, *nous*⁴¹ *et vous*) ont fait couler beaucoup d'encre non seulement dans le domaine des expressions pronominales, mais aussi dans la littérature des expressions déictiques. Dans les approches classiques, on se contente de définir ces pronoms comme des formes déictiques « dont le référent est identifié à partir de la situation de discours où ils sont employés » (Riegel *et al.*, 2018 : 363). Cependant, ils méritent une description plus précise. Nous suivons donc l'approche de Charolles (2002) afin de présenter les caractéristiques, et surtout les fonctions

⁴¹ Du point de vue sémantique, nous pensons que, selon le contexte, « on » (lorsqu'il est employé afin de désigner « nous ») qui est un pronom de troisième personne peut également être compris dans cette catégorie.

référentielles de ces pronoms.

Généralement, dans l'optique référentielle, les pronoms de première et de seconde personne sont appelés « les pronoms de dialogue » parce que, comme nous venons de le voir dans la définition classique, ils renvoient exclusivement aux participants du dialogue présents dans la situation d'énonciation. Les pronoms personnels de ce type ont donc un caractère déictique. Nous pouvons donc dire que ces pronoms n'ont pas de référents constants. Ce renvoi déictique n'est valable que dans le contexte d'occurrence parce que ce sont des pronoms dont la référence ne peut être établie que par un retour sur la situation dans laquelle ils sont produits : en dehors de cette situation spécifique, les référents de *je*, *tu*, *nous* et *vous* ne sont plus les mêmes. Dans ce sens, Charolles (2002 : 195) souligne qu'« ils sont spécifiquement dédiés à la référence aux locuteurs et interlocuteurs engagés dans un échange ».

3.2.1.1 Le pronom *je*

Du point de vue classique, on dit qu'il s'agit d'une expression sui-référentielle renvoyant à l'individu qui dit *je*. D'après Benveniste (1966 : 252), le pronom *je* renvoie à « l'individu qui énonce la présente instance de discours contenant l'instance linguistique *je* ». Dans cette perspective, Reboul (1991 : 128) introduit le point de vue de Milner (1976) sur les pronoms de première et de seconde personne : ils ont une référence virtuelle mais pas d'autonomie référentielle. Autrement dit, ces pronoms personnels qui nécessitent un « complément fourni par chaque énoncé singulier » dans lequel ils se situent ne sont pas autonomes, mais possèdent un certain sens lexical. Pour Milner, le sens lexical de *je* est « celui qui peut dire *je* ».

En revanche, Charolles (2002 : 196-207) contredit cette proposition concernant le caractère du *je* comme suit :

- « *je* » n'est pas équivalent à la description définie *la personne qui prononce « je »*,
- « *je* » ne réfère pas à une instance de discours mais à la personne qui l'énonce,
- « *je* » ne renvoie pas nécessairement à une personne (à savoir, une entité humaine).

Dans cette partie, nous évoquerons les arguments de l'auteur sur le caractère de *je* puis décrirons la manière dont sa fonction référentielle est réalisée.

De prime abord, Charolles (2002 : 196-198) insiste sur le fait que le pronom de première

personne du singulier n'est pas toujours remplaçable par la description définie *la personne qui prononce « je »* :

(31) a. **J'**adore le chien

b. **La personne qui prononce je** adore le chien.

En apparence, nous avons l'impression que *je* de (31a) et *la personne qui prononce « je »* de (31b) sont équivalents. En effet, la différence entre le pronom personnel et la description définie est subtile. Cette différence est plus visible dans l'exemple suivant :

(32) a. **Je** finirai mes études.

b. **La personne qui prononce je** finira ses études.

Lorsque le pronom de première personne *je* de (32a) est remplacé par *la personne qui prononce « je »* de (32b), les deux expressions ne sont pas synonymes. Le référent de (32a) est le locuteur, la personne qui est en train de le dire alors que dans (32b), il y a une possibilité de renvoyer à toutes les personnes qui l'énonceraient.

Deuxièmement, contrairement à la définition traditionnelle qui considère que le pronom de première personne *je* renvoie à une instance de discours, Charolles (2002 : 198-199) souligne que *je* renferme seulement le sujet de l'énonciation mais qu'« il s'expose en tant qu'ego, en tant que moi ». L'auteur cite Kleiber (1986 : 16) qui met également en avant que, dans la phrase « je suis fou », « mon intention première n'est pas de signifier que 'la personne qui prononce cette occurrence de *je*' est folle, mais tout simplement d'indiquer que MOI, je le suis ». Cela suggère qu'une expression déictique ne peut pas être considérée comme « étant une propriété que le référent est seul à posséder » contrairement à la description définie.

En dernier lieu, le pronom *je* n'est pas forcément réservé à la personne. Charolles (2002 : 199-201) mentionne la possibilité dont il peut réaliser une même fonction que le SN démonstratif. Voici deux phrases proposées par l'auteur dans le cadre de la comparaison :

(33) a. **Je** rends la monnaie.

b. **Cet appareil** rend la monnaie.

Imaginons qu'il y a un distributeur de boisson avec une affiche. Le pronom *je* dans (33a) renvoie à la machine et fonctionne donc de manière équivalente au SN « cet appareil » (Il en va de même pour une affiche d'avertissement ou des devinettes). Ainsi, l'emploi de ce type du pronom *je* qui réfère à une « chose » non-animée attend une action ou une réaction d'une façon ou d'une autre du destinataire potentiel. Ces observations montrent que ce pronom

renvoie à son référent d'une manière plus directe que ne le font les autres types d'expressions. Cette idée sur la référence directe insiste sur le fait que « le pronom de première personne adhère d'une manière très étroite à la situation dans laquelle il est énoncé » (*ibid* : 201). Sous cet angle, nous mettons souvent le pronom *je* sur le même plan que deux autres éléments de l'instant de l'énonciation, le *ici* et *maintenant*. Pourtant, *je* a un statut un peu particulier. Tandis qu'*ici* et *maintenant* « repèrent l'état de choses auquel réfère l'énoncé », *je* ne renvoie pas à « une dimension intrinsèque aux états de choses » mais plutôt à un actant de la prédication, comme l'explique Charolles (2002) :

Une chose cruciale qui est que *je*, au niveau le plus primitif, sélectionne non pas une personne déterminée mais un actant manifestant qu'il veut échanger quelque chose avec un destinataire et qui attend de lui une réaction. A ce stade, *je* signale simplement qu'il y a une entité quelconque qui se trouve en connexion actancielle avec l'interlocuteur. (Charolles, 2002 : 203)

3.2.1.2 Le pronom *tu*

Traditionnellement, on définit le pronom de seconde personne singulier *tu* (*vous*)⁴² comme « celui à qui on peut dire *tu* » ou « la personne à qui je parle ». Ce pronom obtiendrait sa valeur référentielle en rapport avec *je*. Autrement dit, comme le signale Benveniste (1966 : 253), le pronom *tu* a une définition symétrique avec celle de *je* : *tu* se définit comme « l'individu allocuté dans la présente instance de discours contenant l'instance linguistique *tu* ». Au sein d'une conversation, *tu* désigne le destinataire alors devenu l'interlocuteur sélectionné par *je*. Les pronoms *je* et *tu* sont donc liés par une relation réciproque : *je* peut devenir *tu* et *vice-versa* dans une allocution. Cependant, Reboul (1991) montre que cette définition classique peut être infirmée par une proposition comme « si je ne te disais pas *tu*, si je ne parlais à personne, tu n'existerais pas ».

Charolles (2002 : 207-210) présente les différentes fonctions référentielles de *tu* : il peut être utilisé de plusieurs façons différentes. *Tu* s'emploie pour se référer à un animal ou à une chose, en d'autres termes, à un référent non-humain, en lui conférant le statut de personne. L'auteur présente un exemple de ce type : un marin parle devant l'océan qui fait rage, « *Tu* ne m'auras pas encore cette fois ! ».

⁴² Le choix entre *tu* et *vous* (*vous* singulier qui est la forme respectueuse de *tu*) peut fonctionner comme un indice pour définir l'interlocuteur.

Ensuite, lors d'un emploi écrit, le pronom *tu* ne désigne pas toujours le lecteur. Par exemple, en lisant le titre d'un livre « Seras-tu là ? », le lecteur a conscience que le *tu* employé ici ne renvoie pas forcément à lui-même mais à une tierce personne. Son interprétation est faite en fonction du contexte. Le lecteur se doit de trouver le bon référent chaque fois où il rencontre *tu* ou *vous* (en tant que pluriel de *tu*) lors de sa lecture.

Les pronoms de seconde personne peuvent s'employer aussi « attributivement » pour emprunter le terme de Charolles (*ibid* : 209) : ils peuvent référer à tout ce qui pourrait être concerné. C'est bien le cas de l'affiche indicative figurant sur un lavabo, « Vous pouvez vous laver les mains ici ». Dans ce cas-là, *vous* renvoie à toute personne susceptible de lire cette annonce et qui a la volonté de se laver les mains.

A partir de ces observations, Charolles (2002 : 238) insiste sur le fait que les pronoms de seconde personne sont moins directs que *je* : « les pronoms de seconde personne, surtout au pluriel, offrent une image dégradée de la saisie indexicale directe dont la forme la plus parfaite est *je* ».

3.2.1.3 Les pronoms *nous* et *vous*

Selon Benveniste (1966 : 233), le pronom *nous* doit être considéré comme « une *jonction* entre '*je*' et le '*non-je*' » : il ne s'agit pas d'« une multiplication d'objets identiques ». L'auteur insiste sur le fait que ce qui constitue le *nous*, c'est la présence du *je* et qu'« il n'y a de '*nous*' qu'à partir de '*je*' ». Ainsi, le *nous* est un pronom qui renvoie à un ensemble comprenant le *je*.

Vous désigne aussi un ensemble dans lequel le *je* est exclu : le pronom *vous* comprend le *tu* et réfère à un ensemble d'interlocuteurs.

D'après Charolles (2002 : 211), les emplois de *nous* et *vous* peuvent se caractériser de manière généralisante ou pluralisante, c'est-à-dire qu'en mentionnant ces pronoms, le locuteur veut soit parler du comportement collectif de l'ensemble qu'il vise, soit pluraliser le sujet.

Ces deux pronoms sont souvent plus complexes qu'ils n'y paraissent parce qu'il n'est pas toujours facile de préciser qui est inclus dans le groupe désigné. Les phrases de l'exemple (34) ci-dessous font référence à une situation d'énonciation dans laquelle le locuteur confirme le lieu de rendez-vous à son interlocuteur par téléphone :

- (34) a. **Je** pars de l'Avenue des Vosges. **Tu** peux me rejoindre à la place Kléber dans 10 minutes.

3.2.2 Les pronoms de troisième personne

Les pronoms de troisième personne (rang 3 et 6⁴³) ont un statut particulier. Ils réfèrent, contrairement à ceux de première et de deuxième personne, à des êtres qui se placent à l'extérieur du dialogue, autrement dit, aux non participants de la communication. Benveniste (1966 : 230) souligne trois points sur les caractéristiques de *il* en le comparant avec *je* et *tu* :

En effet une caractéristique des personnes 'je' et 'tu' est leur *unicité* spécifique : le 'je' qui énonce, le 'tu' auquel 'je' s'adresse sont chaque fois uniques. Mais 'il' peut être une infinité de sujets – ou aucun. [...]

Une seconde caractéristique est que 'je' et 'tu' sont inversibles : celui que 'je' définis par 'tu' se pense et peut s'inverser en 'je', et 'je' (moi) devient un 'tu'. Aucune relation pareille n'est possible entre l'une de ces deux personnes et 'il', puisque 'il' en soi ne désigne spécifiquement rien ni personne.

Enfin on doit prendre pleinement conscience de cette particularité que la 'troisième personne' est la seule par laquelle une *chose* est prédiquée verbalement. (Benveniste, 1966 : 230)

Les pronoms de troisième personne sont spécialisés pour renvoyer aux référents soit humains soit non humains qui sont au centre de l'attention des interlocuteurs. Contrairement à *je* et *tu*, ils peuvent désigner tous les référents, même ceux qui se trouvent hors de la situation de communication. Ils peuvent fonctionner non seulement de manière déictique mais aussi de manière anaphorique. A partir de ces différences, Maillard (1974 : 61) catégorise les pronoms de la personne en deux sous-catégories : les co-locuteurs qui comprennent *je* et *tu* et les non-locuteurs comme *il*. Cette catégorisation a pour but d'éviter l'emploi du terme « personne » qui risque d'être ambigu et de rebaptiser ce terme « catégorie de la locution ». En conséquence, selon l'auteur, le pronom personnel *il* est aussi appelé « délocutif » alors que les pronoms *je* et *tu* sont respectivement « locuteur » et « allocuteur ». Mais ce point de vue est parfois contesté. En effet, Kleiber (2017) souligne que cette définition est incomplète parce qu'elle s'appuie sur le fait que ces pronoms se placent à des niveaux linguistiques

⁴³ Charolles (2002 : 184-185) souligne que l'appellation « pronoms de troisièmes personnes » est morphologiquement et sémantiquement inapproprié parce que ces pronoms ne sont pas pratiquement réservés aux seuls humains. Ils se prêtent à des emplois personnels et non personnels : dans la phrase « Il a plus de dix ans », « il » peut désigner un enfant, un chien, un meuble, etc. Cependant, dans notre recherche, nous les appelons « pronoms de troisièmes personnes » pour des raisons de commodité.

différents⁴⁴. D'après lui, *je* et *tu* possèdent un double rôle « qui combinent rôle d'énonciation et rôle d'énoncé ». Cet aspect qui semble raisonnable, ne nous empêche pas pourtant d'avancer dans cette partie : notre intérêt se porte sur les pronoms de troisième personne qui se placent certainement sur l'instance d'énoncé.

Charolles (2002) mentionne cette caractéristique des pronoms de troisième personne et leur façon de rester attachés au contexte d'émission :

[...] le référent du pronom de troisième personne n'émerge pas de l'énonciation comme cela est le cas avec *je*, il n'est pas lié à l'interaction comme c'est le cas de *tu/vous* qui désigne un allocutaire, il ne fait pas non plus allusion à un ensemble contextuellement agrégé à cet allocutaire, mais à une entité humaine ou non qui est au centre de l'expérience que les participants à la communication peuvent avoir d'un événement notoire concomitant à l'énonciation. (Charolles, 2002 : 217)

C'est donc à partir d'« un événement saillant » que les interlocuteurs atteignent l'entité visée par le pronom. Imaginons que le locuteur voie arriver une collègue qu'il n'attendait pas et qu'il s'adresse à un autre collègue :

(35) a. Mais qu'est-ce qu'**elle** fait là ?

b. Ah oui. **Elle** habite dans ce coin.

Le pronom *elle* réfère directement à la collègue perçue parce que cette situation est au centre de l'attention des interlocuteurs. Autrement dit, il désigne le référent visé en situation immédiate et ce référent est un actant majeur impliqué dans la situation en question.

Charolles mentionne également les emplois inférentiels des pronoms personnels de troisième personne. Ils peuvent référer de façon indirecte par le biais d'indices concernant la situation d'énonciation :

Les pronoms indirects exploitent des traces d'un événement impliquant les destinataires et auquel le locuteur s'associe. (Charolles, 2002 : 229)

Imaginons qu'il soit minuit et que nous entendions un bruit assourdissant provenant de

⁴⁴ « La raison de cette incomplétude est que les trois définitions, malgré leur habit formel ' locutoire' commun, ne relèvent pas du même niveau. Les définitions des deux premières personnes (*je* et *tu*) se placent en effet au niveau de l'énonciation, alors que celle de troisième (*il*) est valide au niveau de l'énoncé. *Je* et *tu* sont des instances d'énonciation, c'est-à-dire se voient attribués un rôle dans l'énonciation, alors qu'*il* est donné comme une instance d'énoncé, c'est-à-dire se voit attribué un rôle dans l'énoncé. » (Kleiber, 2017 : 159).

l'appartement d'à côté. Nous pouvons dire « Ils font encore la fête ». Ici, même si le référent n'est pas physiquement présent durant la situation d'énonciation, nous comprenons que *ils* renvoie à des voisins qui font la fête grâce à l'ensemble des indices collectés dans la situation donnée. Les pronoms de ce type peuvent aussi s'employer métonymiquement dans certaines circonstances. Charolles (*ibid* : 231) propose une situation dans la cuisine où le locuteur a un couteau à la main et demande à son interlocuteur : « Où est-ce que tu les ranges ? ». C'est le cas où un exemplaire conduit, par le pronom, à la classe entière à laquelle il appartient. Pour que ce type d'emploi soit compris sans problème, il faut un élément déclencheur bien perceptible dans la situation d'énonciation. Sinon, on a généralement tendance à chercher l'entité humaine correspondante au pronom utilisé⁴⁵. Ensuite, nous pouvons également évoquer l'emploi collectif des pronoms personnels de troisième personne. Dans une phrase telle que « À Strasbourg, ils font du vélo pour aller au boulot », le pronom *ils* réfère naturellement au collectif des habitants de Strasbourg. Voici un autre exemple similaire :

(36) Devant **la boutique Chanel chaussures**, une longue queue s'étend sur une vingtaine de mètres. « Je fais la queue depuis presque une demi-heure, j'espère qu'**ils** auront encore le modèle. », s'impatiente Julie.. — (*Libération*, 23/06/2021)⁴⁶

Nous pouvons constater que le référent n'est pas clairement visible. Cet exemple ne pose pas de problème malgré le fait qu'il y a un désaccord du genre et du nombre entre la source et le pronom : *ils* renvoie à un collectif d'individus indéterminé en rapport avec la boutique « la boutique Chanel chaussures ». Les emplois de ce type montrent que c'est dans le cadre de la cohérence contextuelle et sur le plan pragmatique que la référence est réalisée.

Ainsi, contrairement aux pronoms de première et de deuxième personne, ceux de troisième personne du singulier réfèrent de manière indirecte même lors d'un emploi déictique : la récupération d'un nom qui correspond à leur genre et leur nombre est exigée pour interpréter *il* et *elle*⁴⁷. Dans la plupart des cas, le référent auquel ces pronoms renvoient est

⁴⁵ Le fait que ces pronoms renvoient parfois aux entités non-humaines montre pourquoi Charolles (2002) n'utilise pas la dénomination « les pronoms de troisième personne » (voir la note 43 *supra*). Cependant, il souligne qu'il y a la « prédilection » de ces pronoms pour les personnes.

⁴⁶ « Réouverture de la Samaritaine : 'Tout a gardé son charme' »
https://www.liberation.fr/economie/reouverture-de-la-samaritaine-tout-a-garde-son-charme-20210623_SU7VZ7DCDFF6NGMCUGQMC7NDZY/

⁴⁷ Comme mentionné *via* l'exemple (35), nous admettons les cas exceptionnels pour le pronom personnel de la troisième personne au pluriel *ils*.

l'actant majeur du scénario de la situation d'énonciation, autrement dit, le personnage important d'un événement qui est au centre des préoccupations des interlocuteurs. En outre, ils peuvent aussi référer à une entité inférable (non seulement humaine mais aussi non humaine) à partir des indices donnés par le contexte soit linguistique soit pragmatique.

3.3 Les pronoms personnels en coréen

Si nous prenons la catégorie de la personne pour point de repère, le coréen se situe aux antipodes du français. Cela a déjà été remarqué par Benveniste (1966 : 226-227) : l'auteur explique, en reprenant Ramstedt (1939), que, contrairement au français, la langue coréenne n'exprime pas la personne dans les verbes. Au lieu de conjuguer les verbes selon la personne, en coréen, les formes verbales diffèrent en fonction de l'ordre social du locuteur par rapport à son interlocuteur. Autrement dit, cette langue considère la relation entre les interlocuteurs ou celle qu'ils entretiennent avec l'actant majeur de l'événement avant de choisir la bonne forme verbale. Nous pouvons donc dire que la manière dont on représente la catégorie de la personne en coréen est moins directe et moins « automatique ». Il y a toujours un processus d'inférence à partir des informations fournies non seulement par le contexte textuel mais aussi par des circonstances d'énonciation.

Pour revenir à notre sujet, Park H. S. (2017 : 101) souligne quelques disparités entre les pronoms personnels de première et de deuxième personne et ceux de troisième personne, communément observées en français et en coréen. Comme nous l'avons vu plus haut (voir la section 3.2.1), les deux premiers renvoient à des entités humaines (ou au moins à des entités personnifiées) correspondant aux actants dans le contexte d'énonciation et donc ont un fort caractère déictique. En revanche, ceux de troisième personne peuvent référer à des entités humaines ou non-humaines et s'employer soit de manière déictique soit de manière anaphorique. Malgré ces points universels, cette catégorie grammaticale du coréen présente quelques particularités par rapport à celle du français.

Kim K. S. (2010) récapitule les cinq caractéristiques de pronoms personnels coréens qui se différencient bien du français et de l'anglais :

- Les pronoms personnels coréens n'ont pas de morphèmes (formes) représentatifs fixes, notamment ceux de première et de deuxième personnes,
- Ils ne changent pas leur forme : le rôle syntaxique est exprimé non par la forme mais par la particule de cas avec laquelle ils s'emploient,

- La différence distributionnelle entre l'oral et l'écrit est observée en coréen : certains d'entre eux sont réservés à l'écrit,
- Ils sont, globalement, moins utilisés en coréen et souvent exprimés sous la forme de pronom zéro,
- Ils peuvent accompagner l'épithète comme *gwanhyungsa*, l'adjectif, la relative, *etc.* et construire le SN.

Inspirée de cette description, dans les pages suivantes, nous détaillerons les caractéristiques des pronoms personnels coréens avec des exemples précis.

3.3.1 Les emplois des pronoms personnels et leurs synonymes

Contrairement au français qui a des formes fixes pour les pronoms personnels, en coréen, il n'y a pas de morphèmes représentatifs fixes. Autrement dit, plusieurs mots ayant la même fonction se présentent de manière nuancée (*via* la forme honorifique, familier et littéraire) et à l'aide du contexte. Le tableau suivant récapitule ces différentes formes variées⁴⁸ :

Personne	Nombre	Forme basique	Forme honorifique	Autres formes
1 ^e	Sing.	Na	Jeo	Jim, Gwain, Bonin, <i>etc.</i>
	Plur.	Woori	Jeohee	
2 ^e	Sing.	Neo	Dangsin	Janae, Geudae, <i>etc.</i>
	Plur.	Neohee	Dangsin(deul)	Yeoleobun
3 ^e	Sing.	Geu (m) Geunyeo (f)	Dangsin	
	Plur. ⁴⁹	Geu(deul) Geunyeo(deul)		

Tableau 7 Les pronoms personnels en coréen

⁴⁸ Nous présentons les synonymes des pronoms personnels de première et de seconde personne (*na* et *neo*, qui sont les formes basiques équivalentes à *je* et *tu*) proposés par Chang K. H. (2002 : 156) et Nam K. S. *et al.* (2019 : 54). Il n'y a pas de consensus sur la liste des synonymes des pronoms, en d'autres mots, il n'existe pas de liste exhaustive. Pourtant, nous avons décidé de présenter les mots de vocabulaire proposés par ces auteurs du fait qu'ils montrent effectivement la diversité des expressions.

⁴⁹ Il n'y a pas de forme fixe pour le pronom personnel de troisième personne pluriel. La marque de pluralité « -deul » s'ajoute à la fin du pronom de troisième personne singulier.

En coréen, le bon pronom personnel est choisi en fonction de la situation d'énonciation, plus exactement selon la relation qu'entretiennent les interlocuteurs entre eux. Par exemple, le locuteur et l'interlocuteur peuvent tout deux employer *na* (première personne du singulier) et *neo* (seconde personne du singulier) dans une situation neutre, c'est-à-dire au sein d'une interaction où toute hiérarchie sociale est absente. Néanmoins, ces pronoms peuvent aussi apparaître dans une situation où le statut social du locuteur est supérieur à celui de son interlocuteur. Dans le cas de la forme honorifique de première personne du singulier, *jeo*, celle-ci renvoie au locuteur lui-même lorsqu'il s'adresse à un interlocuteur qui a un statut social supérieur ou qui est plus âgé.

Les usages du pronom de deuxième personne du singulier de forme honorifique, *dangsin*, varient selon le contexte. En règle générale, il s'agit d'un terme respectueux qui s'emploie souvent entre les époux. Mais s'il est utilisé en dehors de la relation conjugale, son usage instaure une certaine distance ou atteste d'une certaine hostilité du locuteur envers son interlocuteur. Ce même terme peut être employé en tant que pronom personnel de troisième personne lorsqu'il renvoie à une personne respectable ainsi que les parents, les professeurs, *etc.*

Certains pronoms de première personne proviennent du chinois. Les pronoms *jim* et *gwain* étaient uniquement employés par le roi et ne sont pas utilisés dans la vie quotidienne sauf dans des contextes littéraires ou métaphoriques et *bonin* est un pronom du langage écrit qui se présente dans des discours publics. Les pronoms de deuxième personne, *janae* et *geudae*, sont des termes respectueux utilisés pour s'adresser à un interlocuteur plus jeune. Et, *yeoleobun* est un pronom désignant l'audience dans des discours publics.

A part les pronoms présentés dans le tableau 7 ci-dessus, Chang K. H. (2002 : 157) insiste sur le fait que les termes concernant la profession ou le statut social tels que « patron », « professeur », *etc.*, ou bien ceux qui indiquent la parenté comme « mère », « père », « frère », *etc.*, jouent un rôle similaire aux pronoms personnels de deuxième personne. Cependant, à notre avis, il s'agit d'un problème de fonction vocative et donc, nous ne considérons pas ces appellations comme des pronoms personnels malgré leur fonctionnalité spécifique.

Le pronom personnel de première personne du pluriel, *woori* est souvent utilisé pour désigner un être au singulier, notamment lors de son emploi en tant que déterminant possessif⁵⁰. Par exemple, en coréen, les expressions telles que « ma sœur », « mon pays »,

⁵⁰ À proprement parler, la notion de *déterminant possessif* du français n'existe pas en coréen. Afin d'exprimer une relation possessive en coréen, il suffit d'utiliser soit la particule possessive « -ui » après le mot renvoyant au

« mon mari » sont remplacées par « notre sœur », « notre pays » et « notre mari ». Un tel usage est uniquement valable pour désigner des objets familiers. Il en va de même pour le pronom *neohee* (le pronom de deuxième personne du pluriel) comme nous pouvons le voir dans l'exemple suivant :

(37) a. 우리 엄마가 만든 과자야.

Woori eomma+ga mandeun gwaja+ya.

Notre maman+cas-suj. a fait biscuit+terminaison.

(C'est) le biscuit que notre maman a fait. (C'est le biscuit que ma maman a fait.)

b. 학교 끝나고 너희 집에 들리게.

Hakgyo kkeutna+go **neohee jib+ae** deulleu+l+gae.

École finir+part-conj. **votre maison+cas-adv.** passer par+futur+terminaison.

(Je) vais passer par votre maison après l'école. (Je vais passer par chez toi après l'école).

Quant aux pronoms personnels de troisième personne, Nam Kisim *et al.* (2019 : 56) soulignent que seuls *geu* (masculin) et *geunyeo* (féminin) appartiennent à cette catégorie⁵¹. En revanche, quelques expressions ayant la forme « *jisi gwanhyungsa* + N » sont souvent considérées comme des pronoms personnels. Regardons les deux phrases suivantes qui ont syntaxiquement la même structure :

(38) a. 너도 그분 알지 ?

Neo+do **geubun** al+ji ?

Tu+par-aux ce+personne connaître+terminaison interrogative ?

Tu aussi connais cette personne ? (Toi aussi, tu connais cette personne ?)

b. 너도 그놈 알지 ?

Neo+do **geunom** al+ji ?

Tu+par-aux ce+type connaître+terminaison interrogative ?

Tu aussi connais cette personne ? (Toi aussi, tu connais ce type ?)

possesseur soit la forme « N ou Pro (possesseur) + N (objet) ». En conséquence, le pronom personnel peut fonctionner en tant que tel comme le « déterminant possessif » français.

⁵¹ En effet, ces pronoms apparaissent, dans la langue coréenne, à partir du début du 20^e siècle afin de faciliter la traduction des romans occidentaux. Dans cette perspective, le coréen se classe parmi les langues où le pronom personnel de troisième personne n'est pas beaucoup développé.

Dans (38), « *geubun* » et « *geunom* » en gras sont des mots composés de la forme de « *jisi gwanhyungsa* + Ndép⁵² ». Ce genre d'expressions fonctionnellement équivalent aux pronoms personnels sont encore problématiques dans la linguistique coréenne. Au sein de notre recherche, nous admettrons leur fonctionnalité mais les classerons comme des SN démonstratifs à l'instar de Nam K. S. *et al. (ibid.)* ; dans le cas contraire, nous pourrions avoir à faire face à quelques interrogations comme « Est-ce que *geu gangaji* (*ce chien*) ou *geu daetongryeong* (*ce président*) sont les pronoms personnels de troisième personne ? ».

3.3.2 Pronoms invariables et leurs rôles syntaxiques

Les pronoms personnels en français changent de forme selon le cas (*tu-te-toi*, voir le tableau 6 dans la section 3.1 *supra*) et cela indique leur rôle dans la phrase. En revanche, en coréen, ils sont formellement invariables et la présence de la particule de cas suffit à représenter le rôle syntaxique d'une entité linguistique comme l'illustrent les phrases suivantes :

(39) a. 그가 그녀에게 다가간다.

Geu+ga **geunyeo**+ege dagaga+n+da.

Il+cas-suj. **elle**+cas-adv. se rapprocher de+présent+terminaison.

Il se rapproche d'elle.

b. 그녀가 그에게 다가간다.

Geunyeo+ga geu+ege dagaga+n+da.

Elle+cas-suj. il+cas-adv. se rapprocher de+présent+terminaison.

Elle se rapproche de lui.

Dans (39a) et (39b), *geu* et *geunyeo* sont employés dans un cas comme sujet et dans l'autre en tant qu'objet. Pourtant, ces deux pronoms gardent toujours leur forme malgré leur position : seule la particule de cas qui s'ajoute à la fin change pour indiquer leurs rôles syntaxiques dans la phrase.

3.3.3 Pronoms réservés à l'écrit

En coréen, certains pronoms personnels sont réservés uniquement à l'écrit tandis que le

⁵² Le nom dépendant (voir la note 31 *supra*).

français ne montre pas autant de différence distributionnelle entre l'oral et l'écrit à part au sein de certains cas d'emplois génériques⁵³ et l'usage du pronom personnel *on*.

En premier lieu, comme nous l'avons mentionné plus haut (voir le tableau 7), les pronoms personnels de première personne, *jim* et *gwain* sont réservés au roi. La Corée n'étant plus un pays monarchique, ils sont uniquement présents dans la littérature coréenne ou dans les films historiques.

Ensuite, *geudae*, le pronom personnel de deuxième personne, est aussi littéraire. Il s'emploie notamment dans les lettres, la poésie, les paroles de chansons, les livres ou les films.

En dernier lieu, nous avons déjà remarqué que les pronoms personnels de troisième personne en coréen ne sont pas si développés (voir la note 51 ci-dessus) que ceux de première et de seconde personne : l'absence de diversité au niveau de termes (voir le tableau 7 *supra*) suggère également cet aspect. Les deux pronoms personnels de troisième personne, *geu (il)* et *geunyeo (elle)*, qui commencent à être usités dans la langue coréenne afin de traduire « he » et « she » des romans anglais, sont employés à l'écrit, ainsi que dans la littérature, les titres de chansons, les films, entre autres. Dans la vie quotidienne, lors d'une référence de la troisième personne, le SN démonstratif fait fonction de *geu (il)* et *geunyeo (elle)*.

3.3.4 La fréquence d'emploi des pronoms zéros

En coréen, les pronoms zéros⁵⁴ sont fréquents. À l'écrit, les pronoms qui jouent le rôle d'objet se présentent souvent sous la forme de pronom zéro. À l'oral, ceux qui fonctionnent comme sujet et objet peuvent tous être exprimés également par cette catégorie grammaticale. C'est en fonction des informations fournies par les contextes linguistique et extralinguistique que les locuteurs coréanophones arrivent à communiquer malgré l'utilisation fréquente du pronom zéro. Tandis que le français nécessite son propre sujet à chaque proposition, en coréen, une seule ou plusieurs occurrences du sujet sont possibles dans un paragraphe :

(40) a. **Il** était là, comme le plus naturel des convives, sauf qu'**il** ne mangeait ni (Ø) ne

⁵³ *Tu* et *ils* peuvent être employés comme génériques à l'oral plus qu'à l'écrit.

⁵⁴ Représenté par le signe (Ø) dans nos exemples.

buvait. — (*Le fantôme de l'Opéra*, Gaston Leroux, p.19)

b. 그(ii)는 전혀 먹거나 마시지 않는다는 것만 빼고는, 보통의 다른 참석자들과 그리 다르지 않은 자연스런 태도로 (Ø) 거기 그곳에 앉아 있었다.

(Sauf le fait qu'il ne mangeait ni ne buvait du tout, avec une attitude naturelle, pas trop différente des autres convives, (Ø) est resté assis là à cet endroit⁵⁵)

La première phrase (a) de l'exemple (40) est extrait d'un roman français et la deuxième (b) représente sa traduction en coréen. Alors que le sujet « il » est présent deux fois en français, dans la version coréenne, le traducteur utilise le pronom personnel en tête de phrase et une seule fois. Voici un autre exemple que nous reprenons à Kim K. S. (2010) :

(41) a. Dès **sa** première jeunesse, **son** air extrêmement pensif et **sa** grande pâleur avaient donné l'idée à **son** père qu'il ne vivrait pas, ou qu'il vivrait pour être une charge à **sa** famille. — (*Le rouge et le noir*, Stendhal, p.16).

b. 소렐 노인은 그(ii)가 어릴 때부터 우울하고 창백한 얼굴을 한 것을 보고 이 자식(**ce gamin**)은 자라지 못할 거다, 자란다고 해도 가족에게 고생만 시킬 것이라는 생각을 버리지 못하고 있었다.

(Le vieux Sorel a vu qu'il avait l'air pensif avec (Ø) grande pâleur depuis (Ø) première jeunesse et a pensé que **ce gamin** ne vivrait pas, ou (Ø) ne vivrait pour être une charge à (Ø) famille.)

Contrairement à (41a) où le pronom personnel et le déterminant possessif renvoyant au même référent sont employés sept fois, le référent n'est désigné que deux fois au sein de la traduction (41b) : par le pronom personnel « il » et par le SN « ce gamin ». Cela montre que le pronom zéro se présente plus fréquemment en coréen par rapport au français et que, parfois, au lieu du pronom personnel, le coréen utilise le SN de forme « *jisi gwanhyungsa* + N ».

⁵⁵ Pour des raisons de commodité et d'efficacité, dans les cas où nous comparons la structure globale du texte de chaque langue comme dans les exemples (40) et (41), nous présentons une partie d'un texte français (roman) et sa traduction en coréen (traduite et publiée par les traducteurs) sans translittération. La version traduite en coréen sera encore suivie d'une traduction littéraire en français (traduit par nous) afin de faciliter la comparaison structurale.

3.3.5 Pronoms personnels et épithète

Les pronoms personnels coréens peuvent être employés avec une épithète et former un SN comme illustré dans (42b) et (43b), alors qu'en français, cela n'est pas possible comme nous pouvons le constater dans les exemples (42a) et (43a) ci-dessous :

(42) a. *belle toi

b. 아름다운 너

aleumdaun neo

belle toi (= toi, qui est belle)

(43) a. *elle habitant en Andorre

b. 안도라에 사는 그녀

Andorra+e sa+neun geunyeo

Andorre+cas-adv. habiter+term.progression elle

Elle habitant en Andorre (= elle, qui habite en Andorre)

Toutes sortes d'épithètes peuvent se placer devant le pronom personnel sans causer de problème grammatical. Dans cette optique, le coréen est moins restrictif que le français.

3.4 Bilan

Les pronoms personnels s'emploient à la place de diverses unités linguistiques comme un mot, un SN, une phrase, ou bien un segment linguistique plus large. Dans les pages précédentes, nous avons introduit la définition linguistique des pronoms et leurs emplois.

Nous avons également souligné que cette catégorie grammaticale coréenne manifestait quelques disparités importantes avec celle du français. En coréen, les pronoms personnels de première et de deuxième personne possèdent une diversité lexicale alors que ceux de troisième personne sont moins développés qu'en français et sont souvent réservés à l'écrit. Syntactiquement parlant, le coréen semble être plus « flexible » que le français : les pronoms personnels du coréen sont formellement invariables même lors du changement de place dans la phrase grâce à la particule de cas. Ils peuvent aussi s'accompagner de l'épithète, ce qui n'est pas possible en français.

Comme remarqué ci-dessus, souvent les SN ayant la forme « *jisi gwanhyungsa* + N » sont considérés comme étant au même rang que *geu* et *geunyeo*. Sans arriver à un

consensus absolu au sein de la linguistique coréenne, nous avons décidé de suivre l'idée que seuls ces deux derniers appartiennent à la catégorie grammaticale du pronom personnel de troisième personne.

4 Les syntagmes nominaux

Les SN sont définis comme des « syntagmes nominaux organisés autour d'un substantif noyau, toujours accompagnés d'un déterminant, parfois aussi d'un ou plusieurs adjectif(s), ou d'autres éventuels compléments à valeur qualitative » (Baumer, 2015 : 54-55). Voici quelques exemples de SN de formes diversifiées :

- (44) a. le bateau
b. un restaurant alsacien
c. ce stylo que j'ai acheté en Corée
d. mon manteau préféré

Dans l'exemple (44), les SN variés, qu'ils soient définis, indéfinis, démonstratifs ou possessifs, ont, cependant une régularité formelle. À savoir, la micro-structure des trois composants sont la détermination nominale, la complémentation (facultative) et le substantif tête. Les déterminants expriment les propriétés du référent ainsi que le nombre, le genre et l'actualisation ; les compléments, qui sont des éléments facultatifs, délimitent la dénotation du référent, c'est-à-dire qu'ils permettent de préciser une entité ou un groupe d'entités. Mais celui qui joue le rôle principal, à savoir représenter la nature du référent, est le substantif. Charolles (2002) souligne également que, contrairement aux noms propres et pronoms, les expressions de ce type « comportent un N qui fournit une indication sur la nature du référent » (*ibid* : 75).

Dans cette partie, nous décrivons d'abord à la fois les divers emplois de chaque SN et leurs fonctions référentielles. Ensuite, nous présenterons les caractéristiques des SN anaphoriques coréens qui montrent les différences d'inter-langues, celles dont nous avons parlé dans le chapitre précédent.

4.1 Les expressions nominales définies

4.1.1 Les emplois des SN définis

La *GMF* (2018) souligne que « l'article défini présuppose l'existence et l'unicité : il n'y a pas d'autre(s) référent(s) accessible(s) qui vérifie(nt) la description de la réalité désignée par le GN » (*ibid* : 283). Suivant cette description, les SN définis dont le déterminant est un article défini, peuvent référer soit à un N spécifique, c'est-à-dire à un référent particulier et concrètement identifiable, soit à un N générique dans le sens qu'ils peuvent évoquer une classe d'entités. Regardons l'exemple suivant qui présente ces différents usages :

(45) a. Regarde **le bateau**

b. **Le bateau** consomme plus d'essence que la voiture.

D'après la *GMF* (2018 : 273), les expressions de ce type peuvent, de prime abord, « identifier un référent particulier dans l'univers de discours ». Imaginons que le locuteur soit au bord de l'eau et qu'il parle à son interlocuteur d'un bateau qui passe devant eux : « Regarde le bateau ! ». Dans ce contexte d'énonciation, « le bateau » réfère à celui qui est en train de passer. Par ailleurs, le même SN défini, « le bateau », peut renvoyer également au référent générique comme dans (45b).

4.1.2 Les SN définis complets et incomplets

Comme le montre l'exemple (46), la forme du SN défini peut varier : sa forme minimale, (46a), est constituée d'un déterminant et d'un nom alors que sa forme expansée peut accepter un ou plusieurs éléments comme un adjectif (46b), un syntagme prépositionnel (46c), une relative (46d) et une construction infinitive (46e).

(46) a. le bateau

b. le bateau rouge

c. le bateau pour 10 personnes

d. le bateau que j'ai réservé pour mercredi soir

e. On n'a pas le bateau de faire...

Sous cet angle, l'échelle de l'accessibilité d'Ariel (1990)⁵⁶ distingue les descriptions définies longues des descriptions définies courtes. Dans Charolles (2002), les deux catégories

⁵⁶ Voir la figure 4 de la section 2.1 dans le chapitre 2.

sont également discernées sous les termes de SN définis complets et incomplets. Selon l'auteur, la différence entre les deux catégories s'appuie sur l'autonomie référentielle de l'expression⁵⁷, en d'autres termes, sa capacité à identifier le référent en dehors des informations fournies par son contexte est un facteur distinctif important.

- (47) a. le bâtiment le plus haut à Strasbourg
- b. la maire du 7^e arrondissement de Paris en 2021
- c. l'entrée du magasin
- d. la porte

Les SN définis énumérés dans l'exemple (47) n'ont pas le même niveau d'autonomie référentielle. En effet, le SN défini de (47a) renvoie nécessairement à un bâtiment particulier et unique en raison du superlatif et celui de (47b) désigne aussi une personne précise qui assume la fonction au moment de l'énonciation. En vertu de leurs composants, le déterminant défini et l'adjectif, ces SN définis fournissent l'information sur la nature de leur référent.

En revanche, il se pourrait que les exemples (47c) et (47d) soient suivis de questions telles que « Quel magasin ? » et « Quelle porte ? » : ils nécessitent plus d'informations contextuelles sur « le magasin » et « la porte » en question afin de trouver le bon référent.

Ainsi, une question subsiste sur le mécanisme concernant la référence du SN défini incomplet comme « la porte » de (47d) ou « le bateau » de (46a) : comment peuvent-ils être interprétés à travers la situation d'énonciation sans susciter une interrogation telle que « Quel bâtiment ? » ou « Quel bateau ? » ?

Charolles (2002 : 81) présente la notion de *contexte restreint*, en citant les travaux suivants qui introduisent diverses appellations comme « ensemble partagé » (Hawkins, 1978), « pragmatiquement pertinent » (Hawkins, 1994) ou « relationnel » (Galmiche, 1986), « domaine d'interprétation » (Corblin, 1987) et « circonstances d'évaluation » (Kleiber, 1986). Il s'agit d'une notion concernant la délimitation de contexte qui permet au défini de sélectionner « une entité comme unique de son type » (*ibid.*). Les participants à la communication exploitent ce contexte délimité pour interpréter le SN défini incomplet en question. De cette présupposition découle la notion d'*effet de connivence* (*ibid.*) : le locuteur et son interlocuteur sont capables de récupérer sans aucun problème les raisons pour lesquelles on présume l'unicité du référent en question.

⁵⁷ Voir la section 3.1 dans le chapitre 1.

Etant donné que le référent du SN défini peut être une entité déterminée à travers le contexte, nous pouvons ainsi affirmer ceci, à savoir : certains SN définis s'appliquent à une entité unique et d'autres non. Dans les deux cas, le locuteur et son interlocuteur sont en mesure de trouver le bon référent sans aucune difficulté à l'aide du *contexte restreint*.

4.2 Les expressions nominales indéfinies

4.2.1 Les emplois des SN indéfinis

Les SN indéfinis peuvent s'employer de plusieurs manières différentes. D'abord, nous pouvons distinguer les SN indéfinis référentiels des non référentiels. Ces derniers ne faisant pas l'objet de notre étude, nous traiterons seulement les emplois référentiels. Ceux-ci sont habituellement classés en deux sous-catégories : les usages spécifiques et génériques. Regardons l'exemple (48) :

(48) SN indéfini spécifique : J'ai acheté **un stylo**.

SN indéfini générique : **Un stylo**, ça sert à écrire.

La *GMF* (2018 : 293-295) décrit les diverses interprétations des SN indéfinis spécifiques. Les usages de ce type peuvent affirmer simplement l'existence du référent auquel renvoie le N comme le cas de *J'ai acheté un stylo* de (48), montrer le caractère partitif de N en question (par exemple, *J'ai secoué un cerisier, une cerise est tombée*), identifier le référent susceptible d'exister (*Nous cherchons un médecin pour nous aider*) et présupposer une variation d'identité (*Chaque matin, je croise un chien dans la rue*).

Dans la même perspective et comme nous pouvons le constater sur la figure 7 ci-dessous, Charolles (2002) distingue six sous-classes d'usages spécifiques. L'exemple suivant présente les phrases proposées par l'auteur (*ibid* : 182) :

(49) a. Spécifié : Je lis un article de Robert Jackson sur les indéfinis paru en 1972 dans le numéro 19 de la Revue de Philosophie Expérimentale.

b. Sous-spécifié : Je lis un article.

c. Distributif : Paul a offert une montre à ses neveux.

d. Lié à un univers : J'ai rêvé que je dansais le tango avec une infirmière.

e. Attributif : Trace un trait.

f. Individu quelconque : Un roman policier, on le lit en trois heures.

Le schéma suivant synthétise divers usages des SN indéfinis :

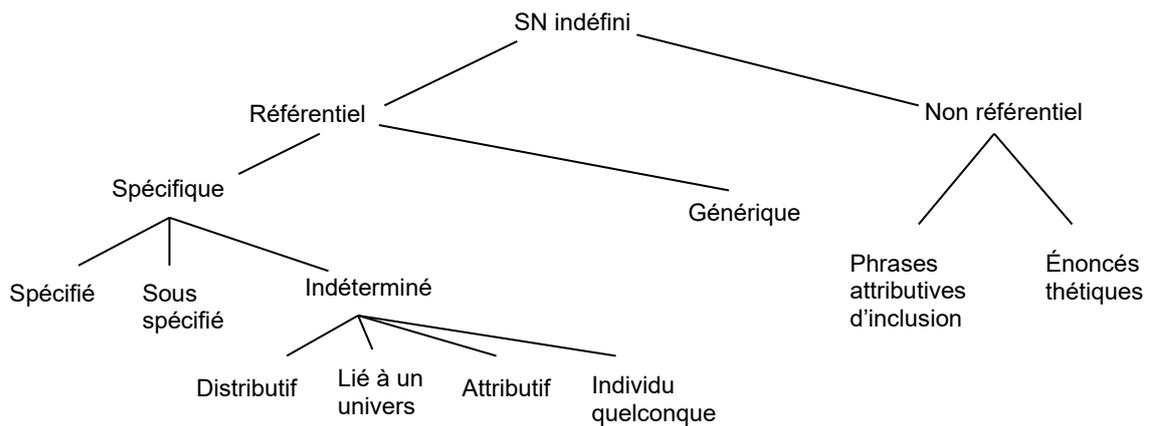


Figure 7 Les emplois des SN indéfinis (Charolles, 2002 : 182)

4.2.2 La procédure d'interprétation des SN indéfinis

Les SN indéfinis ont un point commun avec les SN définis dans le sens qu'ils comprennent le N en tant que tête lexicale du syntagme qui fournit l'information catégorielle sur le référent. Cependant, ils se caractérisent différemment des définis ou des démonstratifs (ce dernier sera traité dans la section 4.3) par leur mode de donation de la référence. Regardons les deux phrases de l'exemple (50) qui pourraient être une réponse de la question « Qu'est-ce que tu as fait ? » :

- (50) a. Je suis allé dans un restaurant alsacien du centre-ville.
 b. Je suis allé dans le restaurant alsacien du centre-ville.

Si la réponse est (50a), l'intention du locuteur est de parler du fait qu'il est allé dans un restaurant spécifique du centre-ville. En revanche, dans (50b), le restaurant en question est considéré comme connu ou mentionné auparavant par les deux participants à la communication ou encore comme étant le seul restaurant alsacien du centre-ville.

Ainsi, l'interprétation des SN indéfinis se réalise d'une façon différente comparée à celle des définis : contrairement au cas d'un SN défini, le locuteur pense que son interlocuteur n'a pas accès au référent auquel la tête lexicale du SN indéfini renvoie. En conséquence, ce qui est exigé pour interpréter le SN indéfini, c'est d'abord les « connaissances lexicales et encyclopédiques générales » (Charolles, 2002 : 143) sur le N concerné. Ensuite, c'est la prédication verbale qui participe à la fixation du référent, c'est-à-dire qu'elle spécifie l'état des choses dans lequel le référent se situe. Charolles (*ibid.*) synthétise les opérations impliquées par l'interprétation des indéfinis chez l'interlocuteur. D'après lui, l'emploi d'un indéfini requiert

de l'interlocuteur :

- a) qu'il instancie (actualise dans son esprit) une classe extensionnelle correspondant au descripteur nominal utilisé,
- b) puis qu'il extraie, sur cette classe, un ressortissant quelconque.

Lorsque cet exemplaire a été actualisé, l'interlocuteur est invité :

- c) à lui attribuer une propriété distinctive dénotée par la prédication verbale principale,
- d) et à réitérer cette procédure autant de fois qu'indiqué par le déterminant.

(Charolles, 2002 : 143)

Cette procédure confirme ce qui est souligné par Ducrot (1970, in Charolles 2002), concernant l'importance de l'énoncé entier dans l'établissement de la référence des SN indéfinis :

Ce n'est pas le SN indéfini qui opère un acte de référence, bien plutôt c'est l'énoncé (...) pris en bloc qui fournit les renseignements nécessaires pour que la référence puisse être opérée ensuite. (Ducrot, 1970 : 156)

4.3 Les expressions nominales démonstratives

4.3.1 Les emplois des SN démonstratifs

Selon la description de la *GMF* (2018 : 286-287), le démonstratif s'emploie de trois manières différentes : déictique, anaphorique et cataphorique (ou mémorielle).

De prime abord, les SN munis de démonstratif peuvent renvoyer immédiatement au référent présent dans la situation d'énonciation. Par exemple, « ce stylo » désigne un stylo qui est accessible, soit physiquement soit contextuellement à partir de la situation d'énonciation.

Ensuite, l'emploi anaphorique des SN démonstratifs présuppose une description préalable qui évoque le référent en question comme « un chien » pour l'exemple (51a) *infra*. Ainsi, cette expression peut être reprise de manière identique ou également par un terme différent⁵⁸ comme un hyperonyme ou un terme évaluatif. C'est le cas pour les exemples (51b)

⁵⁸ Voir la section 2.3.3 du chapitre 1.

et (51c) ci-dessous :

- (51) a. J'ai croisé **un chien** dans la rue. **Ce chien** était trop mignon
- b. J'ai croisé **un chien** dans la rue. **Cet animal** était méchant.
- c. J'ai croisé **un chien** dans la rue. **Ce petit gourmand** voulait mon pain.

Ici, « Cet animal » de (51b) et « Ce petit gourmand » de (51c) sont respectivement un hyperonyme et une expression évaluative relatives au SN « un chien » évoqué dans la phrase précédente.

En dernier lieu, l'emploi cataphorique des SN démonstratifs se réalise lorsque le démonstratif « est justifié par une expansion du N » (*ibid* : 287) qui peut rappeler le référent en question. Reprenons l'exemple (44c), « ce stylo que j'ai acheté en Corée ». Ce SN peut être utilisé de manière cataphorique comme dans « Tu te souviens de ce stylo que j'ai acheté en Corée ? ». En l'occurrence, à la différence de l'emploi anaphorique, le référent n'est pas mentionné dans cette nouvelle situation, mais en vertu du complément employé, le locuteur rappelle un référent déjà connu à son interlocuteur.

4.3.2 L'interprétation référentielle des SN démonstratifs

À la différence du défini qui extrait une entité unique d'une certaine catégorie parmi d'autres de catégories hétérogènes, le démonstratif sélectionne un élément spécifique au sein d'un ensemble d'entités appartenant à une même catégorie, ce qui explique la raison pour laquelle il s'emploie souvent avec un signe d'ostension comme le geste ou le regard. Les deux déterminants se distinguent aussi par le contexte sur lequel ils s'appuient : c'est le contexte matériel de l'énonciation, tel le moment ou le lieu, qui compte dans l'interprétation du démonstratif tandis que l'interprétation du défini se base sur le contexte cognitif, « la représentation des états de choses construites par les destinataires » (Charolles, 2002 : 109) au moment de son occurrence.

Dans cette optique, la référence réalisée par le SN démonstratif, qui « recrute principalement son référent par proximité contextuelle » (Schneidecker, 2021 : 195), est relativement directe par rapport à celle faite par le SN défini et cet aspect du démonstratif est considéré comme la cause de rupture dans la scène en cours :

[...] le SN démonstratif installe le référent dans le focus d'attention des destinataires comme une entité nouvelle ou, en tout cas, en faisant abstraction de la représentation que les destinataires peuvent ou pourraient en avoir dans les circonstances où

intervient l'échange. (Charolles, 2002 : 121)

Cela est évident lorsque nous examinons le cas d'une expression démonstrative « ce dernier »⁵⁹. Dans ses recherches antérieures, Charolles (1995b) souligne que les expressions démonstratives sélectionnant leur référent parmi plusieurs autres issus d'un même contexte éliminent le risque de confusion par rapport au pronom personnel et signalent un certain changement (énonciatif, épistémique, conceptuel⁶⁰).

(52) Après deux ans d'absence, Suzy revient dans l'hôtel où elle avait l'habitude de retrouver André, son amant, le mari de son amie **Cécile**. C'est **cette dernière** qui frappe à la porte. André est mort dans un accident de voiture. Cécile n'a appris leur liaison qu'après le décès. Elle est venue se venger. (Extrait de « La vengeance d'une femme » cité par Charolles, 1995b : 110-111).

Dans l'exemple (52) que nous reprenons à Charolles (*ibid* : 110-111), l'expression « cette dernière » renvoie à celle qui est la plus proche parmi plusieurs candidats à la reprise. À savoir le personnage de « Cécile », qui était moins saillant que celui de « Suzy », subit un changement de statut dans la narration, autrement dit, il est remis dans le focus : les lecteurs,

⁵⁹ « Ce dernier » a un caractère pronominal : il reprend le nom de la personne ou de la chose dont on vient de mentionner. Cependant, nous avons décidé de traiter cette expression démonstrative dans cette partie afin de démontrer le fonctionnement du démonstratif « ce ».

⁶⁰ Schnedecker (2021 : 194-196) souligne que l'emploi de démonstratif peut signaler le changement énonciatif. Par exemple, nous pouvons observer ce fait dans le cas de l'occurrence de « cette dernière » dans (52) qui marque le passage de la fiction à la narration. De plus, son usage peut également évoquer le changement de point de vue, à savoir la façon dont le personnage ou le narrateur perçoit le référent.

ex) Bernis n'a plus que des pensées rudimentaires, les pensées qui dirigent l'action : sortir de ce cirque de montages où la tornade descendante le plonge, où la pluie en rafales est si drue qu'il fait nuit, sauter ce mur, gagner la mer. — (Saint Exupéry, *Courrier Sud*, nous reprenons l'exemple cité par Schnedecker, 2021 : 195).

En dernier lieu, le démonstratif participe à la recatégorisation du référent ou bien à la synthétisation des informations sur le référent fourni par le contexte. Il s'agit du cas d'anaphore résomptive qui récapitule un ensemble d'éléments précédemment évoqués sous une seule expression comme le montre « ce bel édifice » de l'exemple suivant :

ex) La France devait aligner une attaque à faire saliver le monde entier. Avec, comme locomotives, le récent Ballon d'or Karim Benzema et Kylian Mbappé, deux des meilleurs joueurs de la planète, sans oublier le toujours indispensable Antoine Griezmann, et les canonnières Kingsley Coman ou Ousmane Dembélé. Mais patatras, le forfait de Benzema a brisé **ce bel édifice**. — (*Le Monde*, 18/11/2022).

qui suivaient l'histoire de « Suzy », ont, à partir de ce moment, l'impression de lire une histoire insérée entre parenthèses dans la narration, celle de « Cécile », qui est en effet « un retour en arrière sur des événements antérieurs » (*ibid* : 111). En ce sens, comme cité et souligné dans Schnedecker (2021 : 192)⁶¹, le SN démonstratif a « la capacité de marquer une 'rupture' avec l'univers de discours ambiant ». Charolles (1995b : 106-107) explicite également ce mécanisme comme suit :

Le démonstratif, à la différence du défini qui appréhende son référent dans le prolongement des circonstances présidant à son introduction (« le dernier étudiant qu'a reçu Jean »), marque une rupture avec celles-ci. Il se contente d'extraire une entité au sein d'une série de candidats qui sont saisis indépendamment du contexte événementiel.

[...] Le démonstratif, comme beaucoup d'auteurs l'ont relevé, sert à remettre dans le focus un référent qui ne s'y trouvait pas, il permet très facilement soit de repêcher une entité qui était éclipsée par une ou plusieurs autres pour l'installer en position de point de référence, soit d'exhumer un seul candidat parmi deux ou n candidats jouissant d'un même niveau d'attractivité. (Charolles, 1995b : 106-107)

4.4 Les expressions nominales du coréen

Nous avons déjà souligné que le coréen ne possédait pas de système d'article⁶² et, corollairement, qu'il serait insignifiant voire impossible de discerner nettement les SN définis des indéfinis. Comme nous l'avons mentionné dans le chapitre précédent, en coréen, l'ensemble des éléments textuels influence de manière indirecte la détermination d'un SN : la particule, le prédicat et le contexte d'énonciation. Nous avons également parlé de la possibilité que les noms soient déterminés à l'aide des expressions épithètes en citant Hong C. S. (2001). Regardons les trois phrases suivantes :

(53) a. 다섯 아이가 학교 앞에서 실종됐다.

Daseos ai+ga haggyo ap+eseo siljongdwaе+ss+da.

Cinq enfants+cas.suj. école devant+cas.adv disparaître(passif)+passé+term.

⁶¹ Cf. De Mulder (2000) et Charolles (2002).

⁶² Voir la section 2.2 du chapitre 3.

Cinq enfants ont disparus devant l'école.

b. 온 마을이 충격에 빠졌다.

On maeul+i chunggyeog+e ppajyeo+ss+da.

Tout village+cas.suj. choc+cas.adv. tomber(passif)+passé+term.

Tout le village est choqué.

c. 놀란 부모들이 학교에 모였다.

Nollan bumo+deul+eun haggyo+e moyeo+ss+da.

Choqué parents+plur+cas.suj. école+cas.adv. se rassembler+passé+term.

Les parents choqués se sont rassemblés à l'école.

Les trois SN surlignés dans (53) comprennent une épithète devant le nom. Ces expressions ajoutant un sens spécifique comme le nombre, l'intégralité et l'état au référent contribuent à la détermination du nom.

Dans cette partie, nous traiterons brièvement, avec les exemples, des diverses formes des SN déterminés par un élément lexical, à savoir l'épithète, qui se place devant le nom et se classe en cinq sous-catégories selon Hong C. S. (*ibid.*) : *gwanhyungsa*, le possessif, le nom, la construction relative et l'adjectif. Cela permettra de fournir une base de détection des expressions nominales en coréen dans le travail d'annotation.

4.4.1 *Gwanhyungsa* + N

Comme nous l'avons vu précédemment en décrivant les Npr modifiés du coréen, *gwanhyungsa*, qui restreint et qualifie le nom, se divise en trois sous-catégories : *seongsang gwanhyungsa*, *su-gwanhyungsa* et *jisi gwanhyungsa* (voir la section 2.4.1 de ce chapitre).

(54) a. 새 가방

sae gabang

nouveau-ssGHS sac (nouveau sac)

b. 두 가방

du gabang

deux-suGHS sac (deux sacs)

c. 이 가방

i gabang
ce-jisiGHS sac (ce sac)

Seongsang gwanhyungsa représente la nature ou l'état d'une chose comme dans (54a) et *su-gwanhyungsa*, en se plaçant devant le nom comme « du » de (54b), correspond au nombre du référent. Lorsque *jisi gwanhyungsa* est employé, cas de (54c), le référent est désigné de façon déictique comme le fait le démonstratif en français.

4.4.2 La construction possessive + N

Le SN ayant la forme « N1 ou Pro (possesseur) + -ui (particule possessive) + N2 » (voir la note 49 *supra*) exprime la relation possessive comme le montre l'exemple (55) :

(55) 선생님의 책
seonsaengnim+ui chaeg
professeur+cas-possessif livre (livre du professeur)

Les expressions de ce type peuvent également s'employer pour donner d'autres effets sémantiques. Le référent de N1 qui se place devant la particule possessive « -ui », est lié au référent de N2 : il peut être non seulement le possesseur d'une chose comme le montre « professeur » dans (55), mais aussi l'actant d'un comportement (abolement de *chien*), le processus d'un résultat (chemin de *la victoire*), la propriété d'une chose (odeur de *fleurs*), l'apposition (le titre de *président*), voire les divers types de relation métonymique⁶³.

Par ailleurs, la particule « -ui » est susceptible d'être supprimée en fonction des cas. C'est le cas de la forme « N + N » que nous traiterons dans la section suivante.

4.4.3 N + N

Les noms peuvent se placer devant d'autres noms et les qualifier. Les SN de ce type peuvent réaliser divers effets sémantiques. En premier lieu, la relation possessive peut être exprimée par la structure « N1 + N2 » comme le montre l'exemple (56). Il s'agit d'un emploi

⁶³ Les différentes sortes de relation métonymique entre les deux référents sont manifestées par la particule « -ui », indiquant ainsi que le N2, le référent principal du SN, appartient d'une certaine manière au N1. Pour des raisons de commodité et de représentativité définies par le dictionnaire, nous nommons cette structure « la construction possessive + N ».

identique à la forme « N1 ou Pro (possesseur) + -ui (particule possessive) + N2 », ce que nous venons de mentionner dans la section précédente, sauf que la particule possessive « -ui » entre les deux noms est omise : le N1 désigne le possesseur et le N2 renvoie à l'objet :

(56) 선생님 책

seonsaengnim chaeg
professeur livre (livre du professeur)

Le SN de cet exemple désigne donc le même référent que celui de l'exemple (55) *supra* malgré l'absence de particule possessive « -ui ».

Ensuite, le N1 peut fonctionner soit comme un adjectif qui qualifie le N2, soit comme un syntagme prépositionnel. Regardons l'exemple (57) :

(57) a. 기후 변화

gihu byeonhwa
climat changement (changement climatique)

b. 해외 여행

haeoe yeohaeng
étranger voyage (voyage à l'étranger)

Le N1 « gihu (climat) » de (57a) qualifie le N2 « byeonhwa (changement) » et « haeoe (étranger) » de (57b) exerce le rôle du syntagme prépositionnel pour le N2 « yeohaeng (voyage) ».

Cette forme de SN peut s'étendre à plusieurs mots. Comme nous le voyons dans les deux cas de (58) *infra*, l'enchâssement de plusieurs mots est acceptée sans problème :

(58) a. 기차 환승 시간

gicha hwanseung sigan
train correspondance temps (temps de correspondance entre les trains)

b. 대통령 측근 비자금 조성 의혹 관련 조사

daetonglyeong cheuggeun bijageum joseong uihog gwanlyeon josa

Président proche caisse-noire constitution soupçon liaison enquête

(enquête liée au soupçon concernant la constitution d'une « caisse-noire » par les proches du Président)

4.4.4 La construction relative + N

Contrairement à la construction relative en français, la coréenne se place toujours devant le nom pour qualifier ce dernier. Comme le coréen ne possède de pronom relatif, la construction coréenne équivalente est la terminaison « -ㄴ (-n) ». Elle indique que ce qui vient d'être mentionné fonctionne comme une épithète :

(59) 모두가 기다린 영화
moduga gidalin yeonghwa
tout le monde a attendu-**term.** film (film que tout le monde a attendu)

Dans (59), la construction « moduga gidalin (tout le monde a attendu) » est liée au mot « yeonghwa (film) ». La terminaison connective « -ㄴ (-n) » indique la fin d'une proposition, « tout le monde a attendu » et lie cette dernière à l'élément suivant.

4.4.5 Adjectif +N

En coréen, la catégorie adjectivale se subdivise, formellement et syntaxiquement, en deux sous-catégories. Les deux formes de (60) surlignées renvoient à un seul et même adjectif : celui-ci peut se manifester sur le plan verbal comme dans (60a) et sur le plan nominal ce qui est le cas de (60b) ⁶⁴ :

(60) a. 영화가 **좋다**.
yeonghwaga **johda**.
film+cas-suj. bon+term. (film est bon).

b. **좋은** 영화
joheun yeonghwa
bon+term. film (bon film).

Comme nous pouvons le voir dans (60b), le nom « yeonghwa » (film) peut être qualifié par l'adjectif « joheun » (bon+term) qui se place devant lui. Dans ce cas-là, l'adjectif est relié au nom par la terminaison « -ㄴ (-n) ».

⁶⁴ Voir la note 14 du chapitre 3.

4.5 Bilan

Dans cette partie consacrée aux SN, nous avons décrit trois types d'expressions nominales du français : définies, indéfinies et démonstratives.

Ce classement basé sur le caractère déterminatif de l'expression, n'est pas, toutefois, applicable aux SN anaphoriques en coréen puisque la détermination d'une expression référentielle est saisie différemment dans cette langue. Comme explicité dans le chapitre 3, les éléments linguistiques tels que la particule, le prédicat et même le contexte d'énonciation peuvent influencer indirectement la détermination du SN. Par ailleurs, certaines épithètes se manifestant devant le nom peuvent également participer à la détermination d'une expression. Dans la dernière partie de ce chapitre, nous avons donc évoqué les différentes formes de SN anaphoriques coréens : la qualification d'un nom est réalisée par *gwanhyungsa*, par le N employé avec ou sans la particule possessive, par la construction relative ou par l'adjectif.

Durant notre recherche, nous avons pu remarquer que le coréen n'avait pas besoin de la catégorie grammaticale du déterminant possessif ou du pronom relatif pour constituer un SN. Une possession est parfaitement exprimée à l'aide de la seule particule possessive « ui » et l'absence de pronom relatif est compensée par la terminaison connective « -ㄴ (-n) ».

Deuxième partie :
Méthodologie et conventions d'annotation

Chapitre 5.

Constitution du corpus et outils d'analyse

Introduction

Notre étude envisage une analyse des CR à l'aide d'un corpus inter-langues sur un genre discursif spécifique. Ce chapitre a pour but de présenter le positionnement méthodologique que nous avons adopté, ainsi que le corpus sélectionné et les outils utilisés pour mener à bien notre recherche.

Dans la première sous-section (§ 1), nous présenterons d'abord notre corpus composé de deux textes et la raison pour laquelle nous les avons choisis (§ 1.1). Puis, nous détaillerons les caractéristiques et proposerons le résumé de chaque texte, l'un écrit en français et l'autre en coréen (§ 1.2). En outre, nous expliciterons la structure générale du roman avec un schéma narratif et tenterons de l'appliquer à nos textes afin de fournir les informations essentielles sur le corpus (§ 1.3).

La deuxième sous-section sera consacrée à la description du travail d'annotation (§ 2). Nous soulignerons d'abord, sa fonction dans les recherches sur les CR et son processus (§ 2.1). Pour faciliter ce processus, un outil d'annotation sera exploité (§ 2.2) et nous introduirons brièvement les différentes étapes de l'annotation suivies avec SACR (§ 2.3).

En dernier lieu, nous présenterons le logiciel CR Viewer, outil d'analyse des données fournies par l'annotation (§ 3). Dans un premier temps, nous présenterons l'interface du logiciel afin de faciliter la compréhension générale de son fonctionnement (§ 3.1), puis, dans un second temps, nous mettrons en avant les résultats obtenus *via* son utilisation (§ 3.2). En raison d'un caractère relativement pédagogique, cette dernière sous-section permettra à qui le souhaite de travailler sur les CR et d'aborder le sujet plus facilement.

1 Corpus : choix de textes

Etant donné que le corpus témoigne des usages réels d'un phénomène linguistique, il est important de le construire avec des critères précis. Cette sous-section dédiée à la description de notre corpus explicitera les raisons pour lesquelles nous avons sélectionné tels textes et présentera leurs caractéristiques.

1.1 Choix du genre textuel et la constitution du corpus

Comme le souligne Baumer (2015 : 35), chaque genre textuel exige un usage langagier particulier qui pourra contribuer à la réalisation de sa fonction : « la production et la reconnaissance des énoncés est, en partie, déterminée ou influencée par des paramètres liés à la notion de genre ». Le genre textuel est, comme nous l'avons déjà souligné dans le chapitre 2, l'un des trois facteurs qui a des effets importants sur l'analyse des CR.

Parmi les divers genres littéraires, notre attention s'est principalement portée sur le roman. Par rapport aux autres textes courts analysés dans des recherches antérieures¹ tels que les parties introductrices des encyclopédies, les parties d'*incipit* des romans, les portraits journalistiques, les recettes, *etc.*, les romans sont relativement longs. D'ailleurs, ils comportent des passages dans lesquels les phrases employées sont principalement à dominante « déclarative » en raison des buts communicationnels mais aussi des dialogues et ont donc plus de chances que les textes précités de comporter des phrases interrogatives, des impératives et des exclamatives.

Dans un roman, le narrateur raconte l'histoire de personnages engagés dans des événements centraux, et donc les principaux référents sont, dans la plupart des cas, des humains ou des entités humanisées. Toutefois, ces personnages n'ont pas tous le même statut : certains jouent un rôle plus important que d'autres et les textes sont centrés sur un ou plusieurs de ces référents. Ainsi, ce genre littéraire a une structure interne particulière avec un ou des personnages dominants en tant que référent principal. L'occurrence des référents principaux couvre, en général, tout le texte ou au moins plusieurs chapitres dans des cas où certains personnages importants sont introduits tardivement.

¹ Voir la section 3.1 dans le chapitre 2.

De ce fait, nous avons décidé d'analyser quelques chapitres successifs d'un même roman. Ce choix sera pertinent dans le sens que, contrairement aux textes courts mentionnés *supra* qui limitent souvent le nombre et la nature des référents, ce genre de texte pourrait permettre d'observer les CR à long terme. Ainsi, il fera valoir la particularité du roman et nous aidera aussi à avoir une vue globale sur les CR à long terme.

Notre corpus comprend 2 textes, soit 20 165 mots au total dans la dimension inter-langues : un texte français et un texte coréen extraits respectivement d'un roman écrit au début du 20^e siècle². Chaque texte est composé approximativement de 10 000 mots. Voici un tableau récapitulatif de notre corpus :

	Texte 1	Texte 2
Titre	<i>Le Fantôme de l'Opéra</i>	<i>Arbre à feuillage persistant</i>
Genre du texte	Roman	
Nombre de mots	10 422 mots (les trois premiers chapitres)	9 743 mots (les trois premiers chapitres)
Epoque de rédaction	Début du 20 ^e siècle (1910)	Début du 20 ^e siècle (1935)
Langue de rédaction	Français	Coréen

Tableau 8 Tableau récapitulatif du corpus

En faisant abstraction des contraintes inter-langues et de longueur de texte, les deux textes ont été sélectionnés car ils sont libres de droit et facilement accessibles en ligne.

Bien que le corpus soit relativement court, la représentativité des phénomènes de cette étude n'en sera pas moins pertinente parce que notre travail sur des textes longs aura la valeur d'une étude pilote dans le domaine de l'analyse des CR d'un point de vue contrastif.

1.2 Présentation des textes du corpus

Comme nous venons de le présenter, dans cette thèse, nous analysons deux romans différents : l'un est un roman français écrit par Gaston Leroux et publié en 1910, *Le fantôme de l'Opéra* ; l'autre est un roman coréen qui s'appelle *Arbre à feuillage persistant*, écrit par Shim Hoon et publié en 1935.

Initialement, nous avons choisi un autre texte coréen : « Le paysage sous la neige ». Nous avons pensé que ce roman, publié en 1908, pourrait permettre une comparaison valable

² Les différences de 25 ans entre les deux textes seront explicitées dans la section suivante.

avec « Le fantôme de l'Opéra » puisqu'ils datent de la même époque. Cependant, le coréen utilisé au début du 20^{ème} siècle³ ne coïncide pas avec celui d'aujourd'hui, notamment au niveau de l'emploi des pronoms personnels de la troisième personne⁴ alors que les différences semblent moindres entre le français dit moderne et le français contemporain. Ce n'est qu'à partir des années 1930 que les Coréens ont commencé à employer les pronoms personnels de troisième personne. En considérant leur rôle important dans le phénomène des CR, nous avons décidé de remplacer le texte coréen initialement choisi par celui qui a été publié plus tard pour que la comparaison soit plus juste. À la lumière de cette explication, nous pensons que le décalage de 25 ans entre le roman français et le roman coréen ne devrait pas poser particulièrement de problème.

Dans les deux sections suivantes, nous présenterons rapidement les traits et le résumé de chaque roman de notre corpus. Cela facilitera la compréhension de leur structure textuelle lors de l'analyse dans le chapitre 7.

1.2.1 Présentation du roman français : *Le Fantôme de l'Opéra*

Le Fantôme de l'Opéra (désormais *Fantôme*) est un des récits de Gaston Leroux qui montre un aspect complexe de son époque, à savoir le chaos et « le flux d'énergie » (Husson-Casta, 1997 : 11) reflétant le début du 20^{ème} siècle où la modernité et les nouvelles technologies faisaient leur apparition. Husson-Casta (1997) considère que ce roman appartient au genre « troublant » dans le sens où plusieurs genres romanesques comme le roman noir, le roman fantastique, le roman d'aventures et le roman policier y sont mélangés. Cet aspect du roman a des incidences sur la structure et l'introduction des personnages : certains incidents ne sont pas présentés dans l'ordre chronologique et quelques personnages importants ne sont pas introduits dès le début du roman. Le narrateur est durant la plus grande

3 Le coréen employé entre les années 1890 et 1910 est souvent considéré comme « le coréen moderne », ce qui présuppose qu'il existe une différence entre le coréen de cette époque et celui utilisé auparavant ou par la suite (Heo J. Y., 2017 : 10).

4 Outre ce changement, plusieurs autres évolutions de la langue coréenne sont observées à cette époque : les développements du vocabulaire (comprenant le vocabulaire étranger) et de l'uffixe dérivatif, le changement des terminaisons, le raccourcissement de la longueur des phrases et la simplification des expressions honorifiques. Ces changements découlent d'un bouleversement social durant la période d'ouverture vers le monde (par exemple, la France ou les Etats-Unis) à partir de la fin du 19^{ème} siècle et d'un effort des linguistes coréens désireux de mettre en accord langue parlée et langue écrite. (Hong J. S., 2016).

partie du récit, un simple témoin qui rapporte les faits et les paroles des personnages et s'adresse parfois directement aux lecteurs.

Voici le résumé du roman. Celui-ci permettra de comprendre à la fois le contexte des extraits qui seront présentés comme exemples dans les chapitres 6 et 7, et les référents principaux :

- Résumé : L'Opéra de Paris est hanté par un fantôme. Il s'agit en fait d'un musicien défiguré, Erik, qui porte un masque et vit reclus dans les souterrains du bâtiment. Des événements étranges ont lieu lors d'une soirée de gala organisée pour les anciens directeurs au sein de l'Opéra, une soirée qui coïncide avec le triomphe d'une jeune chanteuse, Christine Daaé, remplaçante d'une diva malade.

Le fantôme, un chanteur d'exception avec un visage hideux, donne des leçons de chant à Christine. Au début, cette jeune chanteuse croit qu'il est l'ange de la musique mais elle est effrayée après avoir vu son vrai visage, ce qui lui fait comprendre que c'est le fameux fantôme de l'Opéra. Elle raconte cette histoire au vicomte Raoul de Chagny, son amant, et ils se préparent secrètement à partir loin du fantôme.

Obsédé par Christine et blessé par sa trahison, le fantôme enlève la jeune femme et l'emprisonne dans son royaume souterrain de l'Opéra. Raoul se lance à la recherche de son amoureuse avec le Persan, personnel de l'Opéra, mais il se trouve en danger à cause des pièges diaboliques créés par le fantôme. Christine décide de se sacrifier pour sauver Raoul. Ce noble sacrifice pousse Erik au repentir et il la laisse enfin partir avec Raoul.

1.2.2 Présentation du roman coréen : *Arbre à feuillage persistant*

Arbre à feuillage persistant (désormais *Arbre*) est un roman publié dans un journal en 1935. Dans la littérature coréenne, on distingue trois types de romans écrits à différentes époques : les romans classiques, les néo-romans⁵ et les romans contemporains. *Arbre* est l'un des premiers romans contemporains, à savoir un roman écrit en alphabet coréen (les romans classiques sont dans la plupart des cas écrits en caractère chinois) de la première

⁵ Le néo-roman (Sinsoseol en coréen) signifie littéralement « le nouveau roman » (mais nous utilisons le terme « néo-roman » pour éviter toute confusion avec le terme « nouveau roman » en français). Ce genre littéraire est né vers la fin du 19^{ème} siècle et a perduré jusqu'au début du 20^{ème} siècle. Les romans de cette époque parlent souvent de la modernisation en Corée. Le néo-roman mélange la structure du roman classique et celle du roman contemporain montrant ainsi son caractère transitoire.

moitié du 20^{ème} siècle. Il est écrit d'un point de vue omniscient comme beaucoup d'autres romans de la même époque.

Le roman décrit la situation chaotique de la Corée au début du 20^{ème} siècle où le pays était occupé par l'armée japonaise et présente l'histoire d'un jeune couple qui participe activement au mouvement d'éclaircissement⁶ qui se passe dans tout le territoire.

• Résumé : YoungShin, une jeune fille militante du mouvement d'éclaircissement des ruraux pauvres, participe à un programme national qui a lieu à Séoul et y rencontre un jeune homme DongHyuk qui est aussi militant pour ce mouvement dans son village. Ils deviennent tout de suite amis et partagent les mêmes idées sur le mouvement et l'avenir du pays. Après le programme, chacun rentre dans sa ville mais ils restent en contact en cachant la sympathie qu'ils ont pour l'un envers l'autre.

Un jour, YoungShin rend visite à DongHyuk dans son village et ils tombent tous deux amoureux l'un de l'autre. Ils prévoient de se marier une fois les projets de leur ville respective finis.

YoungShin se dévoue complètement à sa mission et elle tombe malade tandis que DongHyuk reste emprisonné à la place de son frère qui a provoqué un incendie pour montrer sa colère contre les aristocrates. Quand DongHyuk sort de prison, il découvre que YoungShin est morte. DongHyuk, triste mais vaillant, se jure de continuer le mouvement pour honorer la mémoire de YoungShin.

1.3 Aperçu sur la construction textuelle

Afin de comprendre les particularités structurelles du roman, il serait utile d'étudier la structure interne du récit qui est couramment discutée dans le domaine littéraire. Malgré de nombreuses tentatives pour définir la structure, nous parlerons, dans cette recherche, uniquement de la théorie la plus connue et la plus souvent discutée et l'appliquerons à nos textes.

⁶ Le mouvement d'éclaircissement en Corée est un concept à peu près équivalent de celui du « siècle des Lumières ». De la fin du 19^{ème} siècle jusqu'au début du 20^{ème} siècle, il y a eu de nombreuses mobilisations sociales visant la modernisation du pays. Ce mouvement lancé par les élites a évolué vers une sorte de mouvement de résistance contre l'occupation japonaise au début du 20^{ème} siècle. Des jeunes étudiants de la ville se sont portés volontaires afin d'aider les paysans à la campagne et d'instruire les enfants.

1.3.1 Structure textuelle de roman

D'après Genot & Larivaille (1984), le récit est un enchaînement de fonctions narratives qui sont les :

1. Situation Initiale : ensemble de relations stables.
2. Perturbation : modification d'une au moins de ces relations, qui crée un état de déséquilibre.
3. Transformation : action d'un participant de la situation initiale, consistant en modification de relations locales.
4. Résultat : achèvement de la Transformation, résultat local sur les relations directement concernées.
5. Situation Finale : nouvel ensemble de relations stables.

(Genot & Larivaille, 1984 : 285)

Cette conception s'inspire originellement des études d'un folkloriste russe, Vladimir Propp. Il classifie 31 fonctions constituant le conte dans son ouvrage intitulé *Morphologie du conte* (1928). A partir des années soixante, cette idée est traitée par plusieurs auteurs comme Brémond (1966), Greimas (1966), Genot & Larivaille (1984), entre autres. Le terme désignant chaque fonction étant varié selon l'auteur, l'idée elle-même est considérée comme la structure narrative canonique dans le domaine de la narratologie.

D'après ces auteurs, les fonctions narratives sous-tendent tout récit. Pourtant, ils n'appliquent la conception qu'aux contes et récits mythiques et il se pourrait que cette idée ne soit pas applicable aux romans contemporains. Malgré cela, cette notion mérite d'être mentionnée dans cette recherche parce que les deux romans de notre corpus, publiés au début du 20^{ème} siècle, gardent une structure assez « traditionnelle » : ils peuvent être analysés par ces fonctions narratives.

1.3.2 Schéma de la structure narrative du roman

Nous avons schématisé la structure narrative dans la figure (8) ci-après en remplaçant certains termes : « perturbation » a été remplacé par « complication », « transformation » par « péripéties » et « résultat » par « résolution ». Ce schéma comporte cinq étapes comme suit :

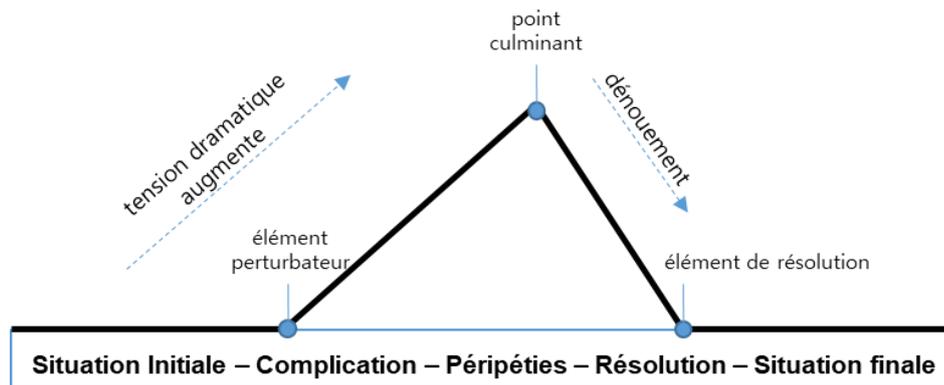


Figure 8 Schéma narratif du roman

Pour commencer, dans la « situation initiale », toutes les informations d'arrière-plan qui fournissent le cadre de l'histoire, le décor, le lieu, et les personnages aux lecteurs sont décrites. Ensuite, la « complication », élément perturbateur, bouscule la stabilité de la situation initiale. Cette étape engendre une série d'événements nommés les « péripéties ». Elle déclenche en même temps les actions des personnages pour éviter ces mêmes péripéties. Au cours de ces événements, la tension dramatique monte jusqu'au point culminant du récit. Les personnages réagissent afin de surmonter la situation complexe de l'apogée, amenant enfin à l'étape de « résolution », la fin des actions. Un nouvel ordre s'instaure dans la « situation finale ».

Inspirée par ce schéma, dans les pages suivantes, nous montrerons la structure et les personnages importants représentés au sein de nos textes.

1.3.3 La structure et les personnages du *Fantôme*

Le roman *Fantôme* est divisé en deux grandes parties et il possède 27 chapitres au total. Chaque chapitre a son propre titre qui résume l'événement central ou représente le personnage qui y joue un rôle important pour donner aux lecteurs un aperçu concernant le contenu à venir. Les trois premiers chapitres qui constituent notre corpus sont intitulés ainsi :

- Chapitre I : Est-ce le fantôme ?
- Chapitre II : La Marguerite nouvelle
- Chapitre III : Où pour la première fois, MM. Debienne et Poligny donnent, en secret, aux nouveaux directeurs de l'Opéra, MM. Armand Moncharmin et Firmin Richard, la véritable et mystérieuse raison de leur départ de l'Académie nationale de musique

Notre travail se limite à ces trois chapitres puisqu'ils suffisent de montrer les CR à long

terme et comme il s'agit du début du roman qui fournit les indices relatifs à l'événement principal du récit et les informations nécessaires pour comprendre les personnages importants. Afin de faciliter la compréhension du déroulement du récit, nous présenterons le schéma narratif de l'œuvre et l'organigramme des personnages entrant en scène durant les trois chapitres de notre corpus :

- Schéma narratif

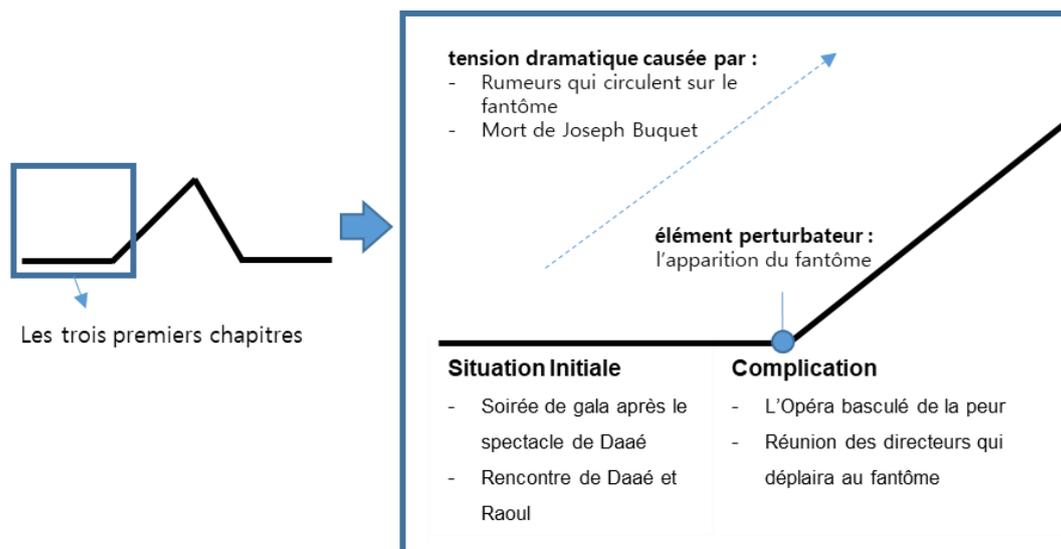


Figure 9 Le schéma narratif des trois premiers chapitres du *Fantôme*

Cette figure illustre la structure de notre texte qui correspond aux deux premières étapes du schéma narratif présenté dans la figure 8 *supra* : la *situation initiale* et une partie de la *complication*. Les deux premiers chapitres fournissent l'arrière-plan du récit en décrivant les événements du passé vécus par des personnages et c'est à partir du troisième chapitre que la *complication* débute.

- Personnages

Etant donné que c'est une histoire qui se déroule à l'Opéra de Paris où, en règle générale, un bon nombre de personnes comme les musiciens, les danseurs, les ouvriers, et les spectateurs se réunissent au quotidien, de nombreux personnages entrent en scène surtout dans les trois premiers chapitres afin de mettre en place le contexte comme le montre la figure suivante :

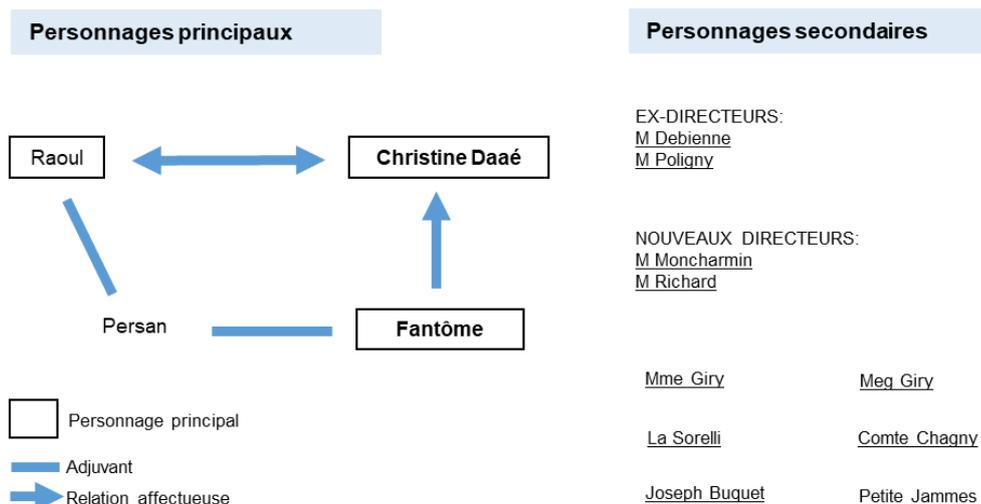


Figure 10 Les personnages dans le *Fantôme*

Nous retiendrons seulement quatre personnages principaux et classerons le reste en tant que personnages secondaires. Raoul et Christine Daaé sont des protagonistes alors que le Fantôme joue un rôle d'antagoniste. Le Persan qui est l'ancien ami du Fantôme raconte l'histoire cachée de celui-ci et aide les deux protagonistes. L'apparition des personnages classés comme secondaires est plutôt épisodique : ils apparaissent majoritairement au début du récit et servent à montrer l'ambiance globale du récit et les caractéristiques des personnages principaux.

1.3.4 La structure et les personnages de l'*Arbre*

Le roman *Arbre* se compose de dix-sept chapitres ; le corpus de cette étude ne traite que les trois premiers de même que dans le roman français. Contrairement au *Fantôme*, les titres de chaque chapitre sont numérotés et ne portent aucune information supplémentaire sur le récit.

Le schéma narratif du texte et la présentation des relations entre les personnages qui suivent nous aideront à contextualiser nos chapitres travaillés :

- Schéma narratif

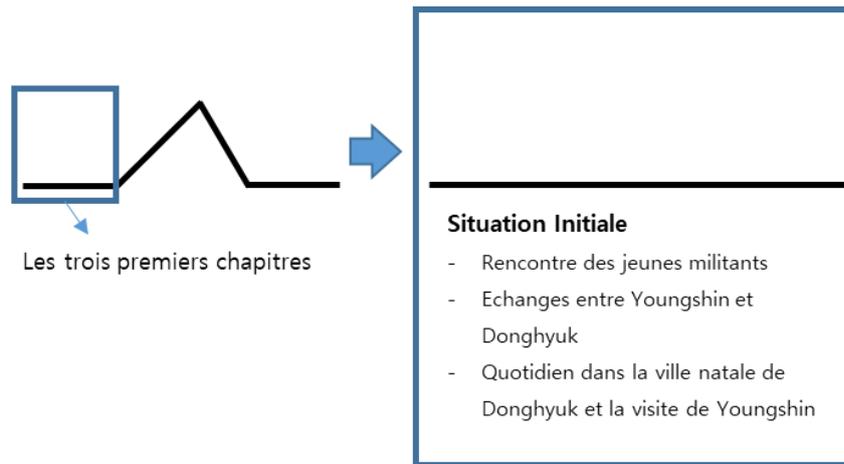


Figure 11 Le schéma narratif des trois premiers chapitres de l'Arbre

A la différence du texte français, les trois premiers chapitres de l'Arbre ne dépassent pas la *situation initiale* du schéma narratif du roman. Chaque chapitre décrit un événement différent comme indiqué dans la figure 11 et les actions des personnages autour de cet événement.

• Personnages

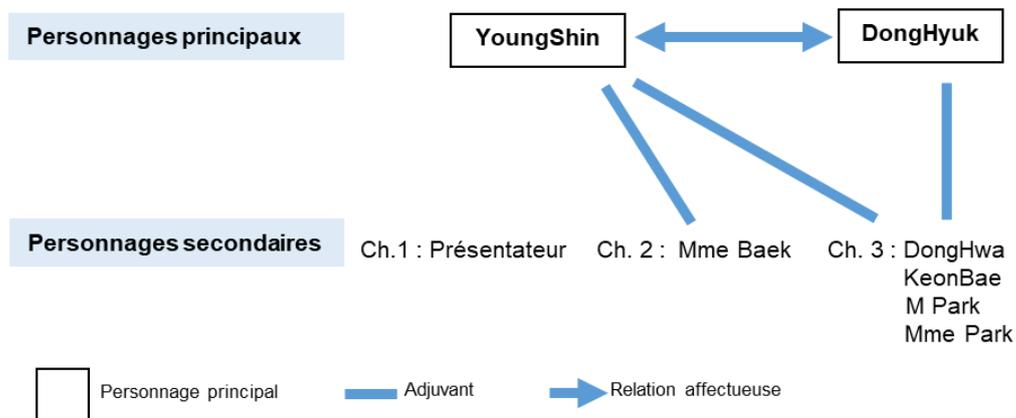


Figure 12 Les personnages dans l'Arbre

Le récit débute par la première rencontre de YoungShin et Donghyuk. Les autres personnages introduits dans les deux premiers chapitres du roman ont un caractère éphémère autrement dit, ils n'apparaissent plus dans les contextes ultérieurs : ils se contentent de jouer un rôle et d'apporter des informations, par le biais des dialogues, sur les deux personnages principaux comme leurs caractères, leurs idées et la situation personnelle dans laquelle ils sont placés. C'est à partir du troisième chapitre du roman que les personnages secondaires établissent un lien assez étroit et constant avec YoungShin et Donghyuk.

Nous venons de présenter nos réflexions méthodologiques quant à la constitution du

corpus et à aux textes choisis. Nous sommes consciente de la relative faiblesse du corpus mais ce choix sera validé, au moins pour une part, par notre objectif de rendre compte des CR à long terme. Dans la dernière partie consacrée à la structure narrative des romans et leurs personnages, nous faciliterons la présentation des analyses à venir. Dans le sous-thème qui suit, nous introduirons le processus d'annotation, le parcours de traitement des textes sélectionnés afin de réaliser l'objectif de cette étude.

2 Travail d'annotation

Le corpus lui-même n'est qu'un ensemble de données brutes. Pour pouvoir exploiter les textes dans notre recherche, nous devons procéder à deux étapes essentielles, l'annotation et l'analyse. La sous-section ci-dessous décrira la première phase de ce processus, l'annotation, en introduisant l'outil choisi pour effectuer la tâche qui lui est attribuée.

2.1 La fonction et le processus de l'annotation

Une fois le corpus constitué, les textes recueillis doivent être annotés afin d'être analysés. Schnedecker *et al.*, (2017) soulignent l'importance de l'annotation comme suit :

L'annotation à la main par des linguistes de phénomènes référentiels et coréférentiels permet, d'une part, de retrouver facilement un exemple illustrant une propriété précise (car la recherche se fait *via* les annotations) et, d'autre part, d'identifier des tendances, voire de commencer à quantifier... (Schnedecker *et al.*, 2017 : 8)

Bien que sa fonction soit claire et utile, l'annotation des CR est souvent considérée comme un processus, physiquement et cognitivement, compliqué et laborieux car elle nécessite des réflexions théoriques sur le problème de la référence. Landragin (2017 : 19) précise les quatre étapes successives de l'annotation et chaque étape a une tâche cognitive à résoudre.

D'abord, l'annotation des CR est sous-tendue par la délimitation des expressions référentielles afin de les repérer dans le texte. Cette étape consiste donc à marquer le point de départ d'une expression référentielle et celui où elle s'arrête. Ensuite, étant donné qu'une chaîne se compose de plusieurs maillons qui renvoient au même référent, il faut résoudre les références dans le texte. Cette deuxième étape concerne les problèmes évoqués par de

nombreux linguistes et philosophes : le problème de l'ambiguïté référentielle et celui des distinctions entre plusieurs types de références (par exemple, générique vs. spécifique ou actuelle vs. virtuelle) qui peuvent affecter l'identification d'un référent. La troisième étape consiste à construire les CR, autrement dit, à regrouper « les expressions référentielles dont les résolutions ont abouti au même résultat, à savoir le même référent » (*ibid.*). Enfin, si elles sont clairement identifiées, nous attribuons à chacune un ensemble de caractéristiques telles que sa place dans la phrase, sa catégorie grammaticale, et son rôle syntaxique. Ces caractéristiques peuvent varier selon l'objectif de l'annotateur.

Ainsi, l'annotation apporte une réflexion sur les divers sujets concernant la référence. Avant d'entamer une discussion concrète sur les problèmes rencontrés au sein de ces étapes et la présentation de nos réflexions pour les résoudre (chapitre 6), nous décrivons, dans la section suivante, l'outil d'annotation et son mode d'emploi.

2.2 Choix d'un outil d'annotation

Compte tenu du caractère complexe des CR et de l'importance de l'annotation, il est pertinent de travailler avec un outil adapté à notre objectif.

Dans le projet DEMOCRAT, afin d'annoter la coréférence, ont été utilisés :

- TXM⁷ : Plateforme d'annotation créée par le projet Textométrie⁸ dans l'intention de fournir un logiciel capable d'aider les utilisateurs à construire et à analyser toutes sortes de corpus textuel numérique.
- SACR (Script d'Annotation des Chaînes de Référence)⁹ : Créé en ligne uniquement pour la coréférence. Cet outil a un format « text-based » qui est simple et à même d'être converti en d'autres formats pour visualiser le résultat.

Les deux logiciels ont respectivement leurs mérites et leurs inconvénients. Etant donné

⁷ <http://textometrie.ens-lyon.fr/>

⁸ La textométrie qui est une notion désignant l'analyse de données textuelles a permis de développer à partir des années 80 des techniques importantes afin d'analyser des grands corpus de textes. Le projet « Textométrie » a pour but de fédérer « les développements logiciels open-source du domaine pour mettre en place une plateforme modulaire appelée TXM ». (<http://textometrie.ens-lyon.fr/>).

⁹ <http://www.boberle.com/projects/sacr/index.html>

que notre corpus comprend le texte coréen qui est complètement différent des textes écrits en alphabet latin, nous avons d'abord testé les deux logiciels sur le plan de l'adéquation avec l'alphabet coréen. SACR étant susceptible de traiter les deux types de token (mot et caractère) s'est révélé plus simple à exploiter avec le coréen, ainsi donc, l'annotation manuelle du corpus est faite à l'aide de l'outil SACR.

Possédant un mode d'annotation intuitif, l'utilisation de SACR facilite une grande partie du travail (Oberlé 2018 : 390-391). Pour créer une chaîne, il suffit d'accomplir trois étapes simples auxquelles tous les utilisateurs sont habitués : cliquer, glisser et déposer. Le logiciel permet non seulement de gagner du temps par l'attribution automatique du nom de référent, c'est-à-dire il propose le nom d'un référent à partir du contenu lexical de l'expression référentielle en question, de diminuer la charge cognitive grâce à la visualisation des maillons, et de chercher plus facilement les maillons et les chaînes spécifiques. Il offre également quelques chiffres clés sur le texte annoté ainsi que le nombre de mots, le nombre de maillons, de chaînes et le nombre moyen de maillons par chaîne. Par ces diverses fonctions, cet outil augmente la vitesse du travail. Oberlé (2018 : 393) souligne que l'annotation par SACR est plus rapide (55 % de temps économisé) comparativement aux autres outils comme TXM ou Analec.

Dans la section suivante, nous présenterons les différentes étapes de l'annotation réalisées *via* SACR.

2.3 Annotation avec SACR

Annoter avec SACR permet d'abord de délimiter les expressions référentielles, de relier de manière visible les coréférences, d'attribuer à chaque maillon un ensemble de caractéristiques, et enfin, de chercher librement celui qui possède telle propriété¹⁰ :

- Délimitation des expressions référentielles : en cliquant le début et la fin d'une expression référentielle, nous pouvons délimiter celle-ci. Elle est encadrée en gris comme dans l'image suivante avec les numéros attribués automatiquement (M1, M2, etc.). L'annotateur peut toujours la renommer. Par exemple, « la partie européenne de la France » peut être renommée « FranceM (France Métropolitaine) » comme le montre la figure 13.

¹⁰ Oberlé (2018), SACR User guide.

Coreference Chain Annotation Tool 5.2.0 ([help](#))

[#1] FranceM La partie européenne de France la France est appelée M3 France métropolitaine et M4 est située à M5 l'extrémité occidentale de M6 l'Europe . M7 Elle est bordée par M8 la mer du Nord M9 au nord , M10 la Manche M11 au nord-ouest , M12 l'océan Atlantique à M13 l'ouest et M14 la mer Méditerranée M15 au sud-est . M16 Elle est frontalière de M17 la Belgique et M18 du Luxembourg M19 au nord-est , de M20 l'Allemagne et de M21 la Suisse à M22 l'est , de M23 l'Italie et de M24 Monaco M25 au sud-est et de M26 l'Espagne et d' M27 Andorre M28 au sud-ouest .

Figure 13 SACR : Délimitation des expressions référentielles

- Regroupement des expressions coréférentielles : celles qui renvoient au même référent peuvent être liées par l'action « glisser-déposer ». Quand les trois expressions référentielles sont regroupées autrement dit, quand une chaîne est construite, les boîtes encadrant ces expressions changent de couleur. Les maillons sont donc encadrés de la même couleur avec la même étiquette comme nous le voyons par « France M » dans l'image ci-dessous. Puis, avec le pop-up, nous pouvons voir la liste des référents (des étiquettes) et de toutes les expressions qui constituent chaque chaîne.

Coreference Chain Annotation Tool 5.2.0 ([help](#))

The screenshot shows the same text as Figure 13, but with boxes around the words. A red box labeled 'FranceM' encloses 'FranceM', 'La partie européenne de France', and 'la France'. A grey box labeled 'France' encloses 'la France'. A pop-up window titled 'Chains and Links' is open over the 'FranceM' box, listing the following referents: 'FranceM', 'La partie européenne de la France', 'France métropolitaine', 'est', 'Elle', and 'Elle'. Below this, a grey box labeled 'France' is partially visible with referents 'la France'.

Figure 14 SACR : Regroupement des coréférences

- Attribution des propriétés : nous pouvons enregistrer le schéma des propriétés dans la page d'accueil avant d'accéder à la page d'annotation. Puis dans la page d'annotation, nous attribuons un ensemble de caractéristiques (catégorie grammaticale, fonction syntaxique et type de référent : ces critères dépendent du schéma enregistré) aux maillons en choisissant parmi les options proposées dans les

menus déroulants (qui montrent ce que nous avons mis dans le schéma) de la barre située en dessous de la page. Le choix d'un schéma peut varier selon nos besoins (Nous aurons l'occasion d'en parler dans le chapitre 6).

Coreference Chain Annotation Tool 5.2.0 ([help](#))



Figure 15 SACR : Attribution des propriétés

- Utilisation de « search box » : SACR nous fournit un « search box » qui est susceptible de trier des maillons possédant certaines caractéristiques. Nous pouvons sélectionner la condition de recherche, par exemple, nous pouvons chercher tous les maillons dont la catégorie grammaticale est Npr, et naviguer entre les maillons concernés.

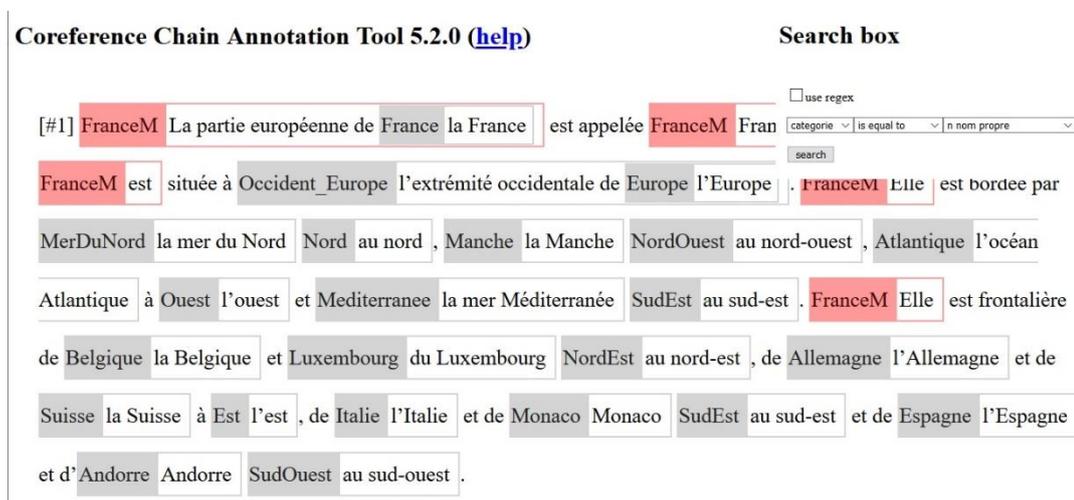


Figure 16 SACR : Utilisation du « Search box »

Les textes annotés *via* SACR seront analysés de manière quantitative et qualitative. Mais avant de passer à l'étape d'analyse, nous soulignerons la nécessité de multiples relectures pour trouver des erreurs qui pourraient être commises et les corriger. Une fois ce travail terminé, nous recourons à un autre logiciel pour analyser les résultats d'annotation.

3 Analyse des données *via* CR Viewer

Les textes annotés sont traités *via* le logiciel, CR Viewer¹¹. Ce logiciel fournit des statistiques relatives aux CR utilisées au fil des textes. À l'aide de ces données quantitatives, nous pouvons procéder à une comparaison inter-langues et voir les différentes caractéristiques des chaînes dans les deux langues. Dans cette sous-section, nous présenterons d'abord l'interface du logiciel avec son mode d'emploi et ensuite, ce que nous pouvons obtenir par l'utilisation de CR Viewer.

3.1 L'interface du logiciel

L'interface principale consiste en deux grandes zones : la gauche montre les options que nous pouvons choisir afin de générer des divers types de données et la zone droite est celle qui affiche des données récupérées. En outre, la partie gauche se divise en deux sous-parties. La première, « select options » (dorénavant, SE-OP), concerne les types de données que nous voulons récupérer et la deuxième, « display options » (dorénavant, DI-OP) se rapporte à la façon soit graphique soit textuelle dont nous les obtiendrons :

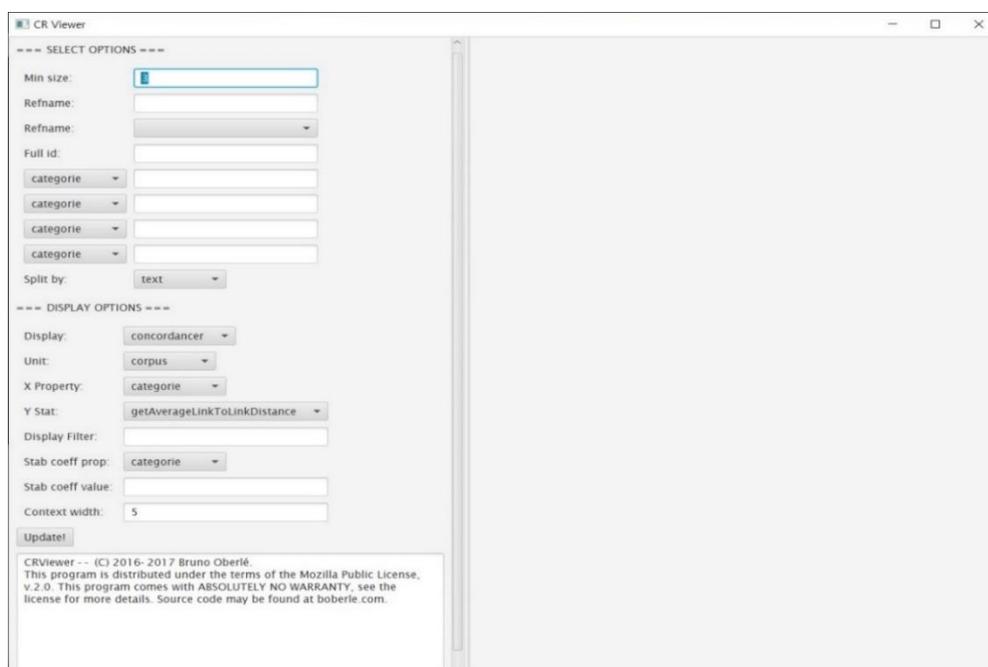


Figure 17 CR Viewer : l'interface principale

¹¹ <https://boberle.com/fr/projects/coreference-analysis-with-crviewer/>

Afin d'afficher le résultat, il suffit de manipuler les deux sections selon nos besoins et d'appuyer sur le bouton « update ».

3.1.1 « Select options » : différents types des données

En enregistrant quelques informations dans la section SE-OP (la figure 18), nous pouvons récupérer les données que nous souhaitons étudier.

=== SELECT OPTIONS ===

① nombre de maillons [Min size: 3]

② nom de chaîne [Refname: []]

③ propriété du maillon [categorie]

Full id: []

[categorie] []

[expansion] []

[fonction] []

[categorie] []

Split by: text

Figure 18 CR Viewer : la section « Select options »

En premier lieu, nous pouvons choisir le type de relation : en fonction du nombre d'expressions référentielles que nous entrons dans « Min size », nous pouvons récupérer tous les référents, les coréférences ou les CR. Pour travailler sur les CR, il faut entrer le chiffre « 3 ».

Toutes les chaînes peuvent être l'objet d'une analyse si nous laissons « Refname » vide. Mais si nous voulons travailler sur une chaîne spécifique, il faut y enregistrer l'étiquette attribuée au cours de l'annotation.

Nous pouvons également décider de trier les expressions référentielles par propriété. Chaque maillon possède certaines propriétés (par exemple, la catégorie grammaticale, la fonction dans la phrase, l'expansion, *etc.*) définies par le schéma d'annotation¹² qui est normalement enregistré dans la page d'accueil de SACR (voir la section 2.3 *supra*). Par exemple, si nous souhaitons trouver tous les maillons dont la catégorie grammaticale est un *pronom personnel*, il suffit de sélectionner « catégorie » dans le menu déroulant de ③ de la

¹² Notre schéma d'annotation sera présenté dans le chapitre suivant avec les conventions.

figure 18 et d'enregistrer dans la case vide à droite le sigle du *pronom personnel* déjà utilisé dans l'annotation.

3.1.2 « Display options » : conditions d'affichage des données

La section DI-OP (figure 19) concerne la façon dont nous pouvons afficher les données. Le logiciel peut énumérer des expressions référentielles, proposer simplement les chiffres clés et présenter soit un histogramme soit un diagramme en camembert.

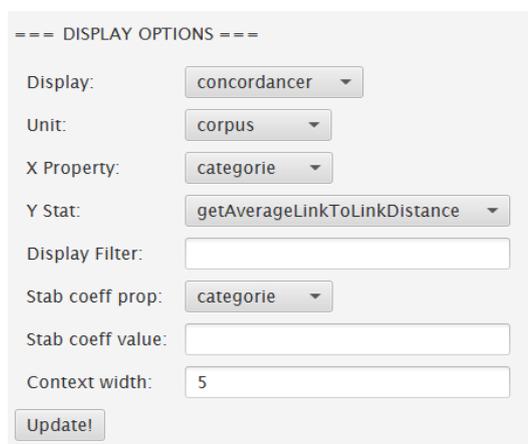


Figure 19 CR Viewer : la section « Display options »

Le menu déroulant « Display » propose cinq options : « concordancer », « stat », « statBars », « frequencis » et « pie ». Parmi ces cinq options, l'option « concordancer » aide à voir tous les maillons annotés au sein de leur contexte d'occurrence. Ils sont énumérés avec le contexte dans lequel ils sont employés. Ensuite, le « stat » présente tous les chiffres concernant les chaînes et le « statBars » visualise ces chiffres (selon notre choix dans le menu déroulant « Y Stat ») sous forme d'histogramme. L'option « frequencis » met en avant les nombres d'occurrences de toutes les propriétés et « pie » représente ces occurrences en diagramme circulaire.

En fonction de l'objectif de recherche, nous pouvons analyser un ou plusieurs textes. Quant aux « statBars » et « frequencis », il est notamment utile d'ouvrir plusieurs fichiers de textes annotés et de les comparer.

3.2 Deux aspects d'analyse

L'analyse de texte annoté par le CR Viewer vise deux objectifs. Le premier est de mettre en évidence la tendance générale du phénomène *via* les chiffres récupérés alors que le

deuxième est d'examiner les caractères individuels des CR.

3.2.1 Tendance globale des chaînes de référence

L'utilisation de CR Viewer permet d'accéder à des chiffres clés susceptibles de décrire la tendance globale des CR apparues tout au long du texte. S'il s'agit d'une recherche comparative où plusieurs textes font l'objet d'une annotation et doivent être analysés ensemble, la récupération de ces chiffres joue un rôle plus important.

Nous avons mentionné plus haut que les chiffres clés sont présentés par le biais de l'option « stats (statistique) » dans le menu déroulant de la section DI-OP. Ce sont des statistiques générales issues de divers paramètres d'analyse concernant les CR. En substance, le logiciel propose douze statistiques : les chiffres généraux (nombre de mots, de paragraphes, de parties, de chaînes et de maillons) et les calculs découlant de ces chiffres généraux¹³ ainsi que la distance moyenne entre les maillons, la longueur moyenne des maillons, la taille moyenne de la chaîne, le coefficient de stabilité, la densité de la chaîne, du maillon et celle de l'annotation. La figure suivante montre une partie de la zone d'affichage présentant le résultat de « stats » :

Id	AvgL2LDist	AvgLinkLength	AvgChainSize	ChainCount	LinkCount	Stab Coeff	ChainDensity
corpus	500.9823529411765	2.3411764705882354	12.423529411764706	170	2112	0.3532566549944273	0.014613599243531334

Figure 20 CR Viewer : l'affichage du résultat de « stats »

Pour plusieurs textes annotés, il est aussi possible d'avoir un histogramme de chaque chiffre et de les comparer avec ceux d'autres textes. Pour ce faire, il faut ouvrir simultanément plusieurs fichiers et choisir « statsBar » dans le menu déroulant de la DI-OP. En choisissant les différents paramètres dans la ligne Y, nous pouvons récupérer les différentes versions des histogrammes. Les trois images présentées dans la figure 21 illustrent un exemple de ce à quoi nous pourrions nous attendre en utilisant l'option « statsBar » :

¹³ Voir le sous-thème 4 du chapitre 2.

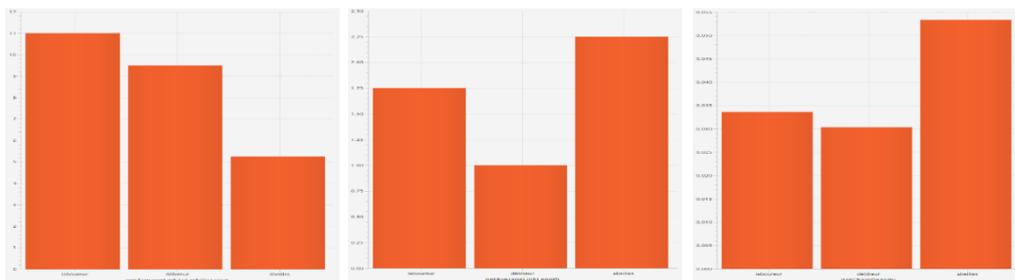


Figure 21 CR Viewer : les trois textes représentés par « statsBar »

3.2.2 Examen du caractère individuel des chaînes de référence

Un autre aspect de l'analyse des CR est d'examiner la composition des chaînes et leurs propriétés. La fonction « frequencies » aboutit au nombre d'occurrences de toutes les propriétés préalablement définies comme dans le schéma suivant :

Id	categorie: i nom indéfini	categorie: s pronom personnel	categorie: z pronom zéro	categorie: p nom possessif	categorie: d nom défini	categorie: f pronom réfléchi	categorie: e déterminant possessif	categorie: r pronom rel
corpus	6	29	8	2	13	8	10	5

Figure 22 CR Viewer : l'affichage du résultat de « frequencies »

La figure 22 montre une partie de l'affichage de « frequencies ». À partir de ce résultat, nous pouvons identifier les fréquences d'apparition de chaque catégorie grammaticale des expressions référentielles employées et leurs fonctions dans la phrase, de chaque type d'expansion, etc. (cela dépend des éléments choisis pour le schéma¹⁴). Cette option permet de rendre compte des propriétés de toutes les CR : en choisissant une ou plusieurs chaîne(s) spécifique(s) dans la section SE-OP et en appliquant « frequencies », nous obtenons les données sur certaines ou toutes CR.

Ces données peuvent également être représentées par un diagramme circulaire comme nous pouvons le voir dans la figure suivante :

¹⁴ Le schéma d'annotation concerne l'attribution des propriétés de chaque maillon (voir la section 2.2 *supra*). Le format de schéma peut varier selon le besoin du chercheur. Nous aurons l'occasion d'en reparler à la fin des conventions d'annotation dans le chapitre 6.

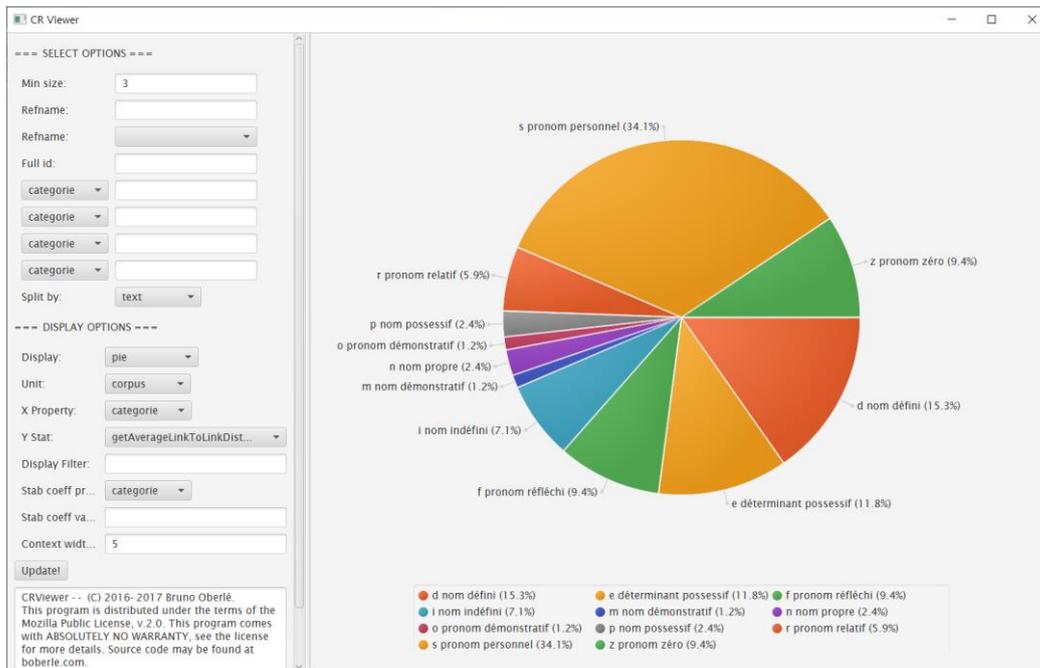


Figure 23 CR Viewer : diagramme circulaire représenté par « pie »

Plusieurs diagrammes sont disponibles en fonction de ce que nous choisissons comme « X Property ». Dans ce menu déroulant, les critères d'analyse que nous avons sélectionnés pour le schéma d'annotation seront présentés.

Afin d'examiner plus précisément chaque expression référentielle, il suffit de choisir la fonction « concordancer ». Tous les maillons concernés sont énumérés avec leurs contextes originels quelle que soit l'option : nous pouvons accéder soit au résultat entier du texte, ce qui est illustré par la figure 24 ci-dessous, soit à celui d'une chaîne particulière.

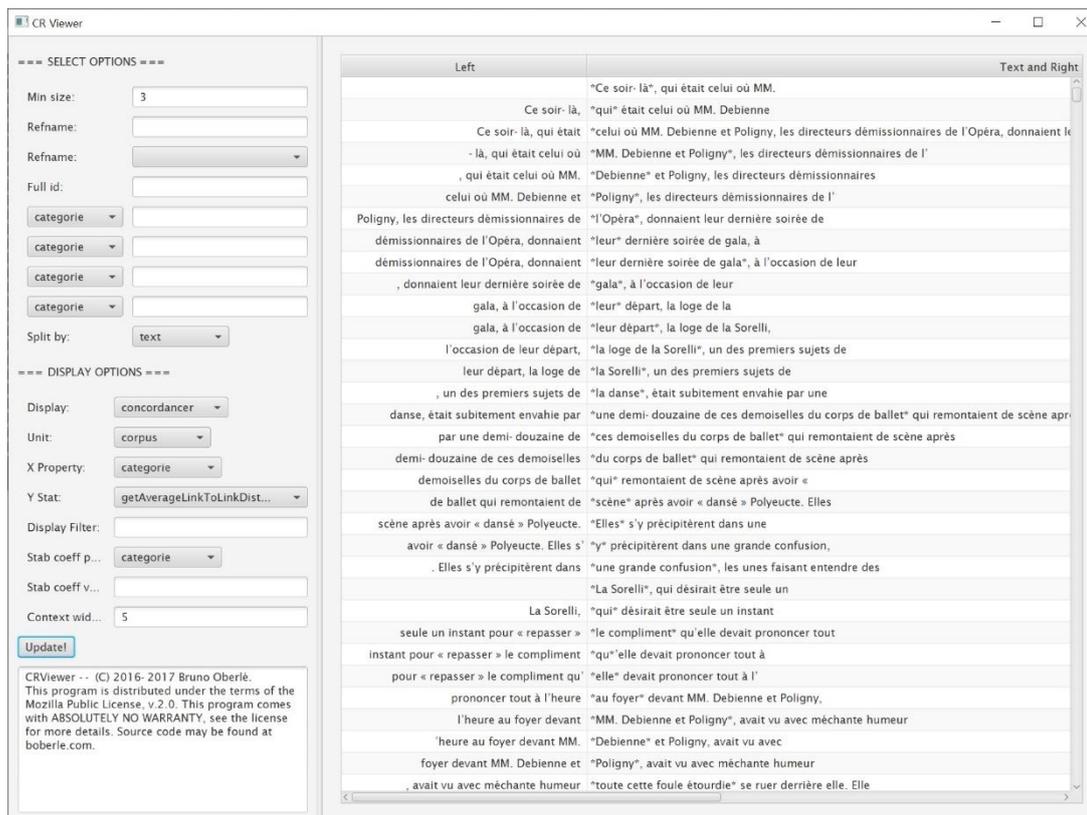


Figure 24 CR Viewer : tous les maillons énumérés par « concordancer »

Comme nous le voyons dans la figure 24, cet affichage est éloigné de l'analyse quantitative. Pourtant il sera quand même utile pour une analyse qualitative. En effet, une telle analyse des données quantitatives peut être détaillée, illustrée, et donc approfondie avec des maillons et des contextes concrets dans lesquels ils sont réellement présentés.

Bilan

Ce chapitre est consacré à la présentation des données de notre corpus, ainsi qu'aux outils choisis pour annoter et analyser les textes. Nous travaillons sur un corpus inter-langues constitué de deux romans de la même époque : les trois premiers chapitres de chaque roman, équivalents respectivement environ à 10 000 mots, font l'objet de notre analyse. Ce choix met en évidence la particularité de notre thèse par rapport aux recherches antérieures de ce domaine qui s'appuient souvent sur des textes relativement courts, permettant ainsi de limiter le nombre et la nature des référents étudiés. Notre thèse se concentrera donc sur une analyse

des CR à long terme dans les deux romans écrits en deux langues différentes.

En outre, nous avons également présenté les outils d'annotation et d'analyse des données ainsi que leurs modes d'emploi. L'annotation exigeant des tâches cognitives sur le problème de la référence est un travail physiquement et intellectuellement fatigant pour les chercheurs. L'utilisation du logiciel SACR facilite, dans une certaine mesure, ce processus d'annotation des textes. Les résultats de l'annotation sont ensuite analysés *via* CR Viewer, un logiciel spécifiquement créé pour les CR, qui permet de rendre compte quantitativement et qualitativement des CR se trouvant dans les textes. Cette dernière partie qui a un caractère « didactique » du fait de l'explication du mode d'emploi de chaque logiciel, pourrait s'avérer utile pour ceux qui veulent travailler sur les CR, notamment en coréen puisqu'il s'agit d'un domaine peu connu dans la linguistique coréenne.

Dans le chapitre suivant, nous présenterons les problèmes rencontrés au sein de l'annotation et nos stratégies pour les résoudre.

Chapitre 6.

Les conventions d'annotation

Introduction

Afin de pouvoir annoter le corpus en vue de l'analyser et d'obtenir des résultats conséquents, il est nécessaire d'établir des critères d'annotation précis. Ce chapitre, dernière étape avant la phase d'analyse en elle-même, y est donc consacré et sera divisé en quatre sous-sections. Notre objectif ici est double : non seulement il s'agit de présenter les critères d'annotation formels et référentiels mais aussi de décrire les résultats escomptés *via* l'annotation.

Dans la première sous-section, nous aborderons les critères de base à respecter lors de l'annotation (§ 1). Comme nous avons déjà présenté les points forts du logiciel et son mode d'emploi lors du précédent chapitre, cette partie se concentrera d'abord sur les points qui n'ont pas été précédemment évoqués, c'est-à-dire la délimitation des expressions (§ 1.1) et le traitement des singletons (§ 1.2). Puis, nous présenterons ensuite le manuel d'annotation du Projet DEMOCRAT et notre position sur ses conventions (§ 1.3).

La deuxième sous-section expliquera nos choix face aux problèmes rencontrés au sein de l'annotation. Ainsi, les points adoptés dans DEMOCRAT seront précisés et ceux nécessitant d'être traités différemment en raison de la disparité inter-langues seront explicités (§ 2).

Nous introduirons ensuite dans une troisième sous-section le schéma d'annotation adopté pour chacun des textes (§3). Ces deux schémas sont construits afin de refléter au mieux les différences inter-langues. Pour ce faire, nous proposerons deux critères communs et deux supplémentaires qui diffèrent selon la langue.

Enfin, tenant compte des conventions d'annotation construites tout au long de ce chapitre, nous mentionnerons brièvement les résultats escomptés de notre étude (§ 4) avant d'analyser les données.

1 Base de démarche d'annotation

Précédemment, nous avons explicité les différentes étapes de l'annotation *via* SACR¹. L'objectif de ce chapitre est de montrer concrètement comment et selon quel principe les expressions référentielles sont délimitées et travaillées. Nous commencerons par mettre en évidence quelques critères de base à respecter lors de l'utilisation de SACR, puis, nous présenterons rapidement le manuel d'annotation du Projet DEMOCRAT², que nous avons adapté aux besoins de notre corpus et de la dimension inter-langues de notre étude.

1.1 Délimiter une expression référentielle

Dans le chapitre précédent, nous avons observé que la commodité de visualisation est l'un des points forts de SACR. À titre d'illustration, examinons l'exemple ci-dessous :

(1)

VincentVanGogh Vincent van Gogh est alors en Provence Provence et, quelques jours plus tard, VincentVanGogh met Desir VincentVanGogh son désir à exécution en peignant, dans RueDArles une rue d'Arles Arles, TerrasseDeCafe VincentVanGogh sa Terrasse de café ...

Détails de l'annotation : Le texte est divisé en segments. Les segments 'VincentVanGogh' sont colorés en rouge et encadrés par des rectangles blancs. Les segments 'Provence' sont colorés en gris et encadrés par des rectangles blancs. Les segments 'Desir' et 'sa' sont colorés en gris et encadrés par des rectangles blancs. Les segments 'Vincent van Gogh', 'son', 'désir', 'une rue d'Arles', 'Arles', 'Terrasse de café' sont colorés en gris et encadrés par des rectangles blancs. Les segments 'RueDArles' et 'TerrasseDeCafe' sont colorés en gris et encadrés par des rectangles blancs.

En délimitant le premier et le dernier mot d'une expression « complexe » comme par exemple *Vincent van Gogh*, les bornes (début et fin) des mentions sont matérialisées par des rectangles avec une étiquette³. Lorsqu'une chaîne est créée, c'est-à-dire lorsqu'au moins trois mentions coréférentielles sont liées, elles sont colorées à l'identique et reçoivent la même étiquette. Il est important de garder le même nom pour le même référent car l'étiquette écrite en majuscules est différente de celle en minuscules. Nous pouvons toutefois attribuer différentes étiquettes, par exemple [Fleur] et [FLEUR], aux mêmes référents dont les caractères identitaires ou référentiels ne sont pas les mêmes⁴. Cependant, dans notre

¹ La section 2.3 du chapitre 5

² Désormais « manuel DEMOCRAT ».

³ Chaque rectangle automatiquement numéroté par SACR peut être renommé par un nom ou un groupe de noms. L'attribution des étiquettes avec une dénomination a pour but de faciliter la reconnaissance de la mention.

⁴ Nous aurons l'occasion d'en parler dans la section 2.1.

recherche, les étiquettes se présentent sous la forme d'un mot doté d'une majuscule qui est la forme basique attribuée par le SACR⁵.

De même, les SN avec des compléments du nom peuvent être traités comme « une rue d'Arles » dans l'exemple (1) où deux expressions référentielles sont imbriquées l'une dans l'autre. Dans de tels cas, nous annotons le complément du nom et le SN entier auquel il appartient : « Arles » et « une rue d'Arles ».

1.2 Traitement des singletons

Les singletons, expressions référentielles qui ne font pas partie d'une chaîne en l'absence de lien coréférentiel avec d'autres mentions, restent dans des cadres gris. Certains chercheurs choisissent d'annoter les singletons⁶ marquant des expressions référentielles isolées en utilisant le code spécial « SI » qui sera ultérieurement remplacé par un identifiant unique de façon automatique⁷. Pour notre part, nous avons exclu les singletons dans l'annotation afin d'alléger le texte et de faciliter la compréhension. Pour ce faire, dès la première annotation, toutes les expressions référentielles sont délimitées étant donné qu'il n'est pas possible de prévoir à l'avance si telle mention constituera ou non une chaîne au sein du texte qui vient. Ensuite, à la relecture, nous excluons des singletons et des coréférences simples (c'est-à-dire les deux mentions coréférentielles qui ne forment pas de chaîne) en supprimant les rectangles qui délimitent ces mentions. Cette stratégie nous aide à rester concentrée à la fois sur notre objectif, les caractéristiques des chaînes, et sur leurs distributions.

1.3 Le manuel DEMOCRAT et notre position

Le manuel DEMOCRAT, constitué de quatre parties (introduction, annotation des expressions référentielles, questions de coréférence et modification de la structure d'annotation avec TXM), fournit une série d'explications sur les questions susceptibles d'être

⁵ Il est possible de changer le nom de mention en appuyant sur les raccourcis claviers « M » et « N ». Ce dernier concerne la numérotation tandis que le premier (« M ») propose le nom de mention.

⁶ Cf. Capin et al. (2021). Cette recherche qui porte sur les textes narratifs du 12^{ème} au 20^{ème} siècle montre qu'entre 70 et 92 % des référents sont des singletons. Nous aurons l'occasion d'en reparler dans le chapitre 7.

⁷ Le manuel DEMOCRAT (2017 : 4).

posées afin de guider les annotateurs à adopter toujours les mêmes procédures.

Notre travail d'annotation se base sur les conventions d'annotation proposées par ce guide, notamment ses deuxième et troisième parties. Le manuel propose d'annoter les noms référentiels (dans certains cas des noms sans déterminants, dans des locutions prépositionnelles non figées, des termes d'adresse, des nombres renvoyant à une date, des attributs nominaux, des mots utilisés en mention ou encore des titres et listes numérotées), les pronoms personnels, démonstratifs, relatifs, adverbiaux (par exemple, *où*) et les indéfinis (*quelques-uns, plusieurs...*), les verbes dans le cas de sujet zéro, le déterminant possessif, les groupes de noms et l'anaphore résomptive. Toutefois, il présente également quelques problèmes de coréférence tels que le traitement de l'anaphore associative, de différents types de référence, de discours direct, des termes équivalents, des référents flous et évolutifs.

Pourtant, comme nous comparons deux langues dont les systèmes sont complètement différents, il se peut que, dans certains cas, un critère soit inutile pour une langue alors que, pour l'autre, il mérite une description. Par exemple, le traitement du pronom relatif doit nécessairement être mentionné pour le français et non pour le coréen dépourvu de cette catégorie grammaticale. De ce fait, nous avons ressenti la nécessité de d'examiner les conventions de DEMOCRAT afin d'atteindre une comparaison pertinente et opératoire.

Dans les pages suivantes, nous parlerons des techniques d'annotation applicables aux deux langues. Les critères nécessitant d'être traités différemment de ce que propose le manuel DEMOCRAT ou ceux qui exigent davantage de débats linguistiques seront justifiés par des exemples issus de notre recherche. Pour des raisons de commodité et de facilité de compréhension, ces exemples seront principalement extraits du corpus français. Cependant, le cas échéant, nous introduirons brièvement quelques extraits du texte coréen afin de démontrer que le phénomène est observé soit d'une manière similaire soit d'une autre (dans le cas contraire, pour être claire, nous mentionnerons que le même phénomène n'existe pas en coréen). Nous espérons ainsi asseoir notre démarche comparative sur des bases solides.

2 Annotation des expressions référentielles

Comme expliqué précédemment, lorsque nous commençons à annoter, nous devons

nous mettre d'accord sur l'annotation des cas discutables. Nous présenterons les problèmes rencontrés durant notre parcours avec des exemples précis, et, proposerons nos propres choix d'annotation, qui pourront être appliqués dans les deux langues.

2.1 Types de référent et de relation coréférentielle

La première question qui se pose est celle de l'objet de l'annotation. Nous pouvons sous-catégoriser les référents selon différents critères : référents humains, non humains (tels que des événements, des lieux, *etc.*), concrets et abstraits, *etc.*

Certains auteurs travaillent exclusivement sur des référents humains (Obry et *al.*, 2017 ; Landragin, 2011 ; Baumer, 2017) comme dans l'exemple qui suit⁸ :

- (2) **[HP_Harry Potter]**, jeune orphelin, a été élevé par **[HP_son]** oncle et **[HP_sa]** tante dans des conditions hostiles. À l'âge de onze ans, un demi-géant nommé Rubeus Hagrid **[HP_lui]** apprend qu'**[HP_il]** possède des pouvoirs magiques et que **[HP_ses]** parents ont été assassinés, des années auparavant, par **[Voldemort_Le mage noir Lord Voldemort]**. **[Voldemort_Voldemort]** avait également essayé de tuer **[HP_Harry]**, mais le sort **[HP_lui]** a rebondi dessus, **[Voldemort_Le]** tuant presque. En fréquentant pour la première fois le monde des sorciers, accompagné par Hagrid, **[HP_Harry]** découvre qu'**[HP_il]** y est très célèbre. **[HP_il]** entame **[HP_sa]** première année d'études à l'école de sorcellerie Poudlard, où **[HP_il]** apprend à maîtriser la magie aux côtés de **[HP_ses]** deux nouveaux amis Ron Weasley et Hermione Granger. Au cours de l'année, le trio se trouve impliqué dans le mystère de la pierre philosophale, gardée au sein de l'école et convoitée par un inconnu qu'ils cherchent à démasquer. L'enquête est menée durant toute l'année jusqu'à la fin où **[HP_le jeune Harry Potter]** rencontre **[Voldemort_Lord Voldemort]**, forcé et contraint de partager **[Voldemort_son]** corps avec celui du Professeur Quirinus Quirell et arrive à le vaincre. — (résumé du film « Harry Potter à l'école des sorciers (2001) »).⁹

Lorsqu'une chaîne est constituée de référents non humains, il pourra s'agir d'un référent

⁸ Les expressions référentielles (comprenant les singletons) qui n'aboutissent pas à constituer une chaîne (son oncle, sa tante, Hagrid, ses parents, des sorciers, *etc.*) ne sont pas marquées dans ce texte. Il en va de même pour les exemples (3) et (4) qui suivent.

⁹ [https://fr.wikipedia.org/wiki/Harry_Potter_\(s%C3%A9rie_de_films\)](https://fr.wikipedia.org/wiki/Harry_Potter_(s%C3%A9rie_de_films))

concret tel que « le sapin » :

- (3) À la différence de chez soi, décorer [**sapin_un sapin de cette envergure**] requiert un savoir-faire particulier et une organisation logistique spéciale. En chiffres, [**sapin_le Grand sapin**] nécessite l'utilisation de deux nacelles de 40m de haut, de grimpeurs communément appelés « nos écureuils », d'une dizaine d'électriciens s'attelant à enrouler les sept kilomètres de guirlandes pour la mise en lumière de [**sapin_l'arbre**] Pour habiller [**sapin_ce sapin**] le massif paysagé se base sur une surface approximative de 200m², réalisé par 8 ouvriers paysagistes. L'ensemble des travaux de décorations représentent trois semaines de travail mais nécessite en amont une longue période de préparation et de réflexion. — (« 3 questions à Romaric Gusto de la société Sigmatech, décorateur du Grand sapin », du site « Strasbourg capitale de Noël ») ¹⁰.

Un référent abstrait ou prédicatif peut aussi devenir l'objet de l'annotation comme le montre l'exemple (4) :

- (4) L'équipe s'est penchée sur [**OptimumHolocène_une période clef de l'histoire climatique de la Terre**], susceptible de nous aider à en anticiper son futur proche, bousculé par nos émissions de gaz à effet de serre. [**OptimumHolocène_Elle**] se déroule il y a 9 000 à 5 000 ans et [**OptimumHolocène_son**] nom de baptême scientifique est « l'optimum Holocène ». Durant [**OptimumHolocène_cette période**], les relations entre le Soleil et la Terre provoquent une insolation particulièrement forte des hautes latitudes nord durant l'été de l'hémisphère nord. — (*Le Monde*, « Sciences² » - blog de Sylvestre Huet, 12/12/2019)¹¹.

Discernant le fait que l'objet de l'annotation, autrement dit, le type de référent que nous annotons, varie selon l'objectif ou le besoin de la recherche, nous avons décidé d'annoter toutes les expressions référentielles qui constituent les chaînes (quel que soit le type de référent) afin d'enrichir la comparaison inter-langues. Cependant, tenant compte de la particularité du genre textuel *roman* où toutes les intrigues se déroulent autour d'un ou plusieurs personnages principaux, nous ferons une double analyse : avec tous les référents dans un premier temps puis seulement avec les personnages principaux dans un second

¹⁰ <https://noel.strasbourg.eu/sous-les-epines-du-grand-sapin>

¹¹ « Climat : une mauvaise nouvelle de 6000 ans ».

<https://www.lemonde.fr/blog/huet/2019/12/12/climat-une-mauvaise-nouvelle-de-6000-ans/>

temps¹². Cette stratégie nous permettra d'avoir à la fois un point de vue global et plus partiel, focalisé sur ce genre textuel.

Nous voudrions ensuite éclaircir notre position sur le problème de l'identité des référents et de la relation qu'ils établissent entre eux. Au cours du texte, des mots ou SN identiques peuvent être utilisés plusieurs fois. Il se peut cependant qu'ils réfèrent à un référent différent. Par exemple, la mention « l'Opéra » peut référer soit au « bâtiment de l'Opéra » soit à la « collectivité de l'institution », ce qui est le cas des exemples (5) et (6) ci-après :

(5) Ce fer à cheval comme toute cette histoire, du reste, – hélas ! – je ne l'ai point inventé, et l'on peut encore aujourd'hui le voir sur la table du vestibule, devant la loge du concierge, quand on entre dans **l'Opéra** par la cour de l'administration. — (corpus « Fantome », paragraphe #29¹³)

(6) Aucun des amis de **l'Opéra** n'ignorait que le cœur de Christine était resté pur comme à quinze ans... — (« Fantome », paragraphe #117)

Ce phénomène référentiel est catégorisé comme « near-identity relation » par Recasens *et al.* (2010)¹⁴. Cette notion peut être définie ainsi : « Si deux SN ne sont ni identiques ni non-identiques, ils sont partiellement identiques comme ils partagent les caractéristiques les plus importantes, mais ils se distinguent au moins une partie essentielle¹⁵ » (Recasens *et al.* 2010 : 151, nous traduisons). La notion de « near-identity » comprend quatre types de relation qui sont les suivantes : la métonymie, la méronymie, la relation entre des entités qui sont dans la même catégorie et enfin la relation des différentes caractéristiques spatio-temporelles d'une même entité. Voici un tableau récapitulatif des sous-types de « near-identity » et leurs exemples¹⁶ :

¹² Nous en reparlerons dans le chapitre 7.

¹³ A partir d'ici, tous les exemples présentés seront des parties extraites de nos corpus. Nos renvois seront réalisés comme suit : chaque exemple sera suivi d'une étiquette indiquant le nom du texte et le numéro du paragraphe dont il est extrait.

¹⁴ Au lieu de s'appuyer sur la distinction binaire, coréférence et non-coréférence, qui ne couvre pas suffisamment tout le phénomène observé dans des textes authentiques, ils définissent la zone du milieu (middle zone en anglais), la zone de « near-identity » afin de catégoriser les entités partiellement coréférentielles.

¹⁵ « If two NPs are neither Identical nor Non-Identical, they are partially the same in that they share most of the important characteristics, but differ in at least one crucial dimension. » (Recasens *et al.*, 2010 : 151).

¹⁶ Pour les noms de 4 types et leurs sous-types, nous utilisons les termes anglais en tant que tels présentés par

4 types de relation de « near-identity »	Définition et sous-types	Exemples
Name metonymy	Une entité peut être désignée <i>via</i> ses diverses facettes. (role, location, organization, information realization, representation, other)	<i>Le Fantôme de l'Opéra</i> est l'un des plus grands succès de l'histoire des comédies musicales. <i>Cette œuvre</i> , publiée en 1910 par Gaston Leroux...
Meronymy	Les composants d'une entité peuvent être utilisés pour désigner l'entité elle-même. (part-whole, stuff-object, set-set)	Jacques Chirac a cumulé <i>les plus hautes fonctions au service de l'Etat</i> . Le <i>Premier ministre</i> , sous la présidence de Valéry Giscard d'Estaing...
Class	Des entités appartenant à la même catégorie avec différentes positions hiérarchiques peuvent être considérées comme une entité « plus spécifique » ou « plus générale ». (more specific, more general)	Comme tous les enfants, je rêvais d'être <i>un super-héros</i> . Ce qui est attirant dans <i>Thor</i> , c'est qu' <i>il</i> est moins connu que les autres.
Spatio-temporal function	Une entité (ou un événement) peut être reconnue dans un lieu ou un temps différents (place, time, numerical function, role function)	<i>La température corporelle moyenne normale chez l'être humain</i> était de 37 °C au 19 ^{ème} siècle, et aujourd'hui <i>elle</i> est de 36,5 °C.

Tableau 9 4 types de relation « near-identity » et leurs sous-types (Recasens et al., 2010)

D'après cette définition, dans nos exemples (5) et (6), les mentions de « l'Opéra » peuvent être considérées comme des éléments de la relation métonymique, et donc ne sont pas discernées. Nous excluons cependant cette notion dans notre travail d'annotation. Au lieu d'annoter les mentions faisant partie de la relation « near-identity » comme une chaîne, nous considérons comme « non-coréférentielles » les entités partiellement coréférentielles et leur donnons les étiquettes avec un indice discernable tel que [Opéra_bâtiment **l'Opéra**] pour le (5) ou [Opéra_institution **l'Opéra**] pour le (6).

les auteurs afin d'éviter l'ambiguïté.

2.2 Les noms et les SN

Nous venons de confirmer que toutes les expressions référentielles participant à la constitution d'une chaîne, quel que soit le type de référent, font l'objet de l'annotation. Par ailleurs, concernant l'identité des référents et la relation coréférentielle que les maillons établissent entre eux, nous décidons de clairement distinguer les référents partiellement identiques et donc de les annoter séparément¹⁷.

Nous allons désormais examiner certains cas qui concernent les SN représentés dans les CR, y compris les SN expansés par des compléments, comme nous l'avons vu précédemment à travers l'exemple (1).

2.2.1 Les expressions figées

Les expressions figées comprenant des SN ou des pronoms qui sont des suites de mots non dissociables entre eux ne font pas l'objet d'annotation. D'après Ruwet (1983 : 34), ce sont des expressions avec un caractère arbitraire, qui sont apprises par cœur :

Les expressions idiomatiques doivent être plus ou moins apprises une à une, par cœur ; le sujet parlant apprenant une langue doit y être exposé, découvrir qu'elle existe dans la langue sous telle forme plutôt que sous telle autre. (Ruwet, 1983 : 34-35)

Ainsi, les expressions de ce type ne sont ni modifiables ni séparables. Gross (1996 : 9-17) souligne qu'elles possèdent certaines caractéristiques, même si la portée et le degré de figement diffèrent selon l'expression : la présence d'une séquence de mots, l'opacité sémantique, le blocage des propriétés transformationnelles, la non-actualisation des éléments, le blocage des paradigmes synonymiques, la non-insertion et la possibilité de défigement pour un effet spécifique. Ces caractères sont souvent synthétisés sous le nom de « non commutation d'un segment » et de « non compositionnalité du sens » qui montrent qu'une rigidité syntaxique et sémantique est une qualité intrinsèque de figement.

Regardons à présent l'exemple (7). L'expressions « des caves au grenier » signifie en fait « tout le bâtiment », autrement dit les éléments de cette expression sont ici employés au sens figuré (ou opaque selon le terme de Gross). Étant donné que cette expression ne fait référence

¹⁷ Par ce choix, nous pouvons déjà exclure les entités constituant une anaphore associative ou le cas des référents évolutifs de l'annotation de notre corpus.

ni aux caves ni au grenier au sens propre, nous ne pouvons substituer par exemple le mot «chambres» au mot «caves». Nous ne pouvons pas non plus insérer d'expansions, notamment un adjectif comme des «des caves sombres aux greniers poussiéreux» sans provoquer de défigement. Enfin, la permutation des SN est impossible, c'est-à-dire qu'on ne peut pas dire «du grenier aux caves» :

(7) Et cela sous peine de devenir la proie de la puissance occulte qui s'était emparée du bâtiment, des caves au grenier ! — (*Fantôme*, paragraphe #28)

Nous n'annotons pas les mots qui font partie d'une expression figée parce que, à titre individuel, ils ne sont pas des référentiels¹⁸.

Le manuel DEMOCRAT considère que certaines locutions prépositionnelles et conjonctives appartiennent également à cette catégorie : *à cause de, à la suite de, à l'exception de, en dépit de, en cas de, etc.* Ce type d'expressions comportant des noms non référentiels, nous suivrons donc ce choix :

(8) Mais vous avez parfaitement raison », acquiesça le docteur, et il mit tout le monde à la porte, à l'exception de Raoul et de la femme de chambre. — (*Fantôme*, paragraphe #145)

Le SN « l'exception » de l'exemple (8) ne peut pas être détaché de la locution prépositionnelle « à l'exception de » et donc n'est pas annotable.

2.2.2 Les noms sans déterminant

En général, les noms employés sans déterminant n'ont pas vocation à référer et donc ne sont pas annotables. Malgré ce fait, le manuel DEMOCRAT conseille de les annoter car ils peuvent être repris. Cependant, comme nous décidons d'annoter uniquement les CR, l'explication du manuel ne suffirait pas à justifier le choix. Dans les pages suivantes, nous présenterons deux aspects de noms utilisés sans déterminant, le cas de référents non

¹⁸ L'exception : si un élément d'une expression figée est repris dans le discours ultérieur et constitue une chaîne, cela devient annotable :

EX) ...on ne saurait dire « nez à nez », car le fantôme n'en avait pas...

Dans cet exemple, l'expression « nez à nez » est figée mais son segment « nez » est repris par « en » dans le contexte ultérieur.

actualisé et celui des noms ayant une fonction classifiante.

2.2.2.1 Référents actualisés et non actualisés

Un nom ne fait pas l'objet d'annotation lorsque son référent n'est pas actualisé¹⁹. Dans l'exemple (9), « gala » n'a pas de référent actualisé tandis que le SN entier « leur dernière soirée de gala » renvoie à une soirée précise. Nous choisissons de ne pas annoter le type d'emploi où un terme est non actualisé :

(9) ... [[leur]_a dernière soirée de **gala**]_b, ...²⁰ — (*Fantôme*, paragraphe #3)

Par ailleurs, un référent non actualisé est parfois repris dans le contexte ultérieur. Dans les exemples suivants, le terme « clef » de (10) est une partie d'une expression figée et donc non actualisé. Ce terme est, pourtant, repris dans (11), le contexte ultérieur, actualisant ainsi « la clef » :

(10) Et elle ferma la porte à [_{Clef} **clef**]. — (*Fantôme*, paragraphe #5)

(11) Mais la Sorelli, armée d'un stylet qui ne la quittait jamais, osa tourner [_{Clef} **la clef**] dans la serrure, et ouvrir la porte, pendant que les danseuses reculaient jusque dans le cabinet de toilette et que Meg Giry soupirait... — (*Fantôme*, paragraphe #37)

Dans de tels cas, nous considérons qu'un lien coréférentiel est établi entre « clef » de (10) et « la clef » de (11). Ainsi, un nom sans déterminant, qui est repris pour établir un lien coréférentiel avec d'autres segments du texte, est annoté.

Un référent actualisé peut parfois porter un nom sans déterminant. Regardons l'exemple suivant :

(12) Du moment qu'un lieutenant de pompiers n'hésitait pas à s'évanouir, [_{CorpsDeBallet} **coryphées et rats**] pouvaient invoquer bien des excuses à la terreur qui [_{CorpsDeBallet} **les**] faisait se sauver de toutes [_{CorpsDeBallet} **leurs**] petites pattes quand [_{CorpsDeBallet} **elles**] passaient devant quelque trou obscur d'un corridor mal éclairé. — (*Fantôme*, paragraphe #27)

¹⁹ Certaines expressions figées peuvent également être incluses dans ce type d'expressions.

²⁰ En raison de facilité de la description, nous utiliserons des crochets afin de matérialiser les bornes des mentions et les faire suivre par une lettre (dans l'ordre alphabétique a, b, c...). Dans le cas où il faut préciser l'étiquette, nous allons la mettre devant le nom comme [_{Clef} **clef**] de l'exemple (8).

Dans l'exemple (12), « coryphées et rats » n'a pas de déterminant, et renvoie en effet à un groupe de gens bien spécifique dans l'Opéra. Cette expression étant, en outre, reprise dans la même phrase par les pronoms personnels, elle est donc annotée.

2.2.2.2 Fonction classifiante

Certains noms utilisés sans déterminant sont équivalents, fonctionnellement, à des syntagmes adjectivaux. Les noms de ce type sont exclus de l'annotation :

(13) ...[le nez cher à **Grévin**]_a , [des yeux de **myosotis**]_b — (*Fantôme*, paragraphe #4)

(14) ...[deux yeux d'**émeraude**]_a — (*Fantôme*, paragraphe #43)

Les usages du même type sont observés sous deux formes différentes lorsque nous annotons le texte coréen. Premièrement, certains des SN dont la forme est « N1 + N2 » peuvent indiquer lorsque le N1 désigne une caractéristique du référent de N2 comme dans l'exemple (15) :

(15) [골키퍼 노릇] — (*Arbre*, paragraphe #220)

Goalkeeper noreut
Gardien de but rôle
le rôle de gardien de but.

Dans tel cas, le N1 ne fait pas l'objet d'annotation parce qu'il renvoie à un référent non actualisé, autrement dit, le N1 représente un aspect caractéristique que possède également le N2. Seul le SN entier sera annoté : [RôleDeGK 골키퍼 노릇 (le rôle de gardien de but)].

Les noms sont ensuite utilisés comme équivalent adjectival lorsqu'ils sont suivis de certaines particules de cas adverbial telles que « -cheoreom » et « -gati ». Ces particules ajoutées à la fin du nom (sujet subsidiaire) indiquent que l'entité postposée (sujet primaire) à ce SN a la même caractéristique que son référent (ou au moins un aspect similaire) comme dans l'exemple suivant :

(16) [황소처럼 건강한 체격] — (*Arbre*, paragraphe #240)

Hwangso+cheoreom geonganghan chegyeog
Taureau+part.ADV robuste+adj. Physique
La physique robuste comme (celle de) taureau.

Dans les constructions de ce type, où deux entités juxtaposées fondent une relation

métaphorique à l'aide de la particule de cas adverbiale, nous n'annotons pas à titre indépendant le SN comprenant cette particule. Sujet subsidiaire de la métaphore, le N de ce SN n'est pas actualisé mais remplit une fonction classifiante comme certains adjectifs. Nous annotons donc la construction entière de la manière suivante : [PhysiqueDeDonghyuk 황소처럼 견강한 체격 (La physique robuste comme celle de taureau)].

2.2.3 La distinction entre référents génériques et spécifiques

Dans cette étude, nous faisons la différence entre les référents utilisés de manière générique et les référents spécifiques comme illustré dans l'exemple ci-dessous où il est bien question de deux fantômes différents : le fantôme en général et le fantôme de l'Opéra.

(17) Et comme elle était la première à croire [Fantome_G **aux fantômes en général**] et à [Fantome *celui de l'Opéra*] en particulier, elle voulut tout de suite être renseignée. [...]

Et c'est vrai que, depuis quelques mois, il n'était question à l'Opéra que de [Fantome *ce fantôme en habit noir*] [Fantome *qui*] se promenait comme une ombre du haut en bas du bâtiment, [Fantome *qui*] n'adressait la parole à personne, à [Fantome *qui*] personne n'osait parler et qui s'évanouissait, du reste, aussitôt qu'on [Fantome *l'*] avait vu, sans qu'on pût savoir par où ni comment. [Fantome *l'*] ne faisait pas de bruit en marchant, ainsi qu'[Fantome *il*] sied à [Fantome_G **un vrai fantôme**]. — (*Fantôme*, paragraphe #7 & #14)

Ce type d'occurrence des référents différents apparaît souvent dans un texte où des séquences descriptives enchâssées construisent un décrochement énonciatif du texte. L'alternance entre le référent générique et le référent spécifique se présente au sein des phrases voisines, des séquences ou des paragraphes différents. À part l'insertion typographique, d'autres dispositifs linguistiques étayent cette rupture : changement du temps verbal (du passé simple et de l'imparfait au présent), emploi des adverbes généralisants (*en général*) et du sujet « on ».

Voici ci-après un autre exemple illustrant également ce type de détachement énonciatif. Nous y trouvons en effet deux types de « lieutenant de pompiers », l'un employé de façon générique et l'autre de façon spécifique :

(18) [Lieutenant_G **Un lieutenant de pompiers**], [Lieutenant_G **c'**]est brave ! [Lieutenant_G **Ça**] ne craint rien, [Lieutenant_G **ça**] ne craint surtout pas le feu !

Eh bien, [Lieutenant *le lieutenant de pompiers en question*], [Lieutenant *qui*] s'en était allé

faire un tour de surveillance dans les dessous et [Lieutenant *qui*] s'était aventuré, paraît-il, un peu plus loin que de coutume, était soudain réapparu sur le plateau, pâle, effaré, tremblant, les yeux hors des orbites, et [Lieutenant *s'était quasi évanoui*] dans les bras de la noble mère de la petite Jammes. Et pourquoi ? Parce qu' [Lieutenant *il*] avait vu s'avancer vers [Lieutenant *lui*], à hauteur de tête, mais sans corps, une tête de feu ! Et je le répète, [Lieutenant_G **un lieutenant de pompiers**], [Lieutenant_G **ça**] ne *crain*t pas le feu.

[Lieutenant *Ce lieutenant de pompiers*] s'appelait [Lieutenant Papin]. — (*Fantôme*, paragraphe #24, #25 et #26)

Les expressions en gras désignent « le lieutenant de pompiers » en général avec un commentaire au présent, autrement dit, les six expressions coréférentielles ne font pas référence à une personne spécifique. Puis le narrateur glisse à un commentaire au plus-que-parfait et les maillons mis en italiques renvoient à un lieutenant de pompiers précis qui s'appelle « Papin ».

Afin de discerner plus facilement le référent générique du référent spécifique dont l'alternance se répète au fil du texte, nous attribuons le code spécial « _G » à la fin de l'étiquette des référents génériques comme « Fantôme_G » et « Lieutenant_G ».

2.2.4 Le dédoublement du collectif

Il convient ici de revoir les divers modes de cohabitation des CR (voir la section 1.2 dans le chapitre 2) : succession, entrecroisement, dérivation, partition/fusion et parallélisme. Parmi ces différents modes de cohabitations, certains cas de dissociation semblent compliqués à annoter. En effet, lorsqu'une chaîne se divise en plusieurs autres chaînes, il est parfois difficile de saisir exactement les référents de ces dernières comme illustré dans l'exemple (19) :

(19) [Danseuses_Loge **Elles**] s'y précipitèrent dans une grande confusion, **les unes** faisant entendre des rires excessifs et peu naturels, et **les autres** des cris de terreur. — (*Fantôme*, paragraphe #3)

Les référents visés respectivement par « les unes » et « les autres » font partie de l'ensemble visé par le pronom « elles » qui renvoie, dans le contexte du roman, à une demi-douzaine de danseuses entrées dans la loge de *la Sorelli*. Cependant, ces SN partitifs

désignent une collectivité indéterminée²¹, à savoir nous ne pouvons pas exactement identifier qui sont « les unes » parmi toutes les danseuses concernées et combien de danseuses sont comprises dans ce groupe. Nous annotons les SN de ce type, seulement s'ils sont repris ultérieurement dans le texte et fondent une coréférence, en leur donnant les étiquettes susceptibles de décrire de quoi il s'agit dans la mesure du possible. Regardons l'exemple (19') :

(19') [Danseuses_Loge **Elles**] s'y précipitèrent dans une grande confusion, [Danseuses_SousGroupe1 **les unes**] faisant entendre des rires excessifs et peu naturels, et [Danseuses_SousGroupe2 **les autres**] des cris de terreur.

Pour « les unes » et « les autres » qui renvoient respectivement aux petits groupes des danseuses dans la loge de *la Sorelli*, nous attribuons les simples étiquettes comme « Danseuses_SousGroupe1 » et « Danseuses_SousGroupe2 ».

2.2.5 Les syntagmes nominaux attributifs

L'attribut est interprété comme un prédicat qui décrit une caractéristique de son sujet soit le sujet de la phrase soit le complément d'objet direct. Dans cette partie, nous parlerons uniquement de la construction attributive du sujet dont l'expression attributive a une forme de « nom sans déterminant » comme dans la phrase « Il est président de la République française depuis 2017 » ou d'un « SN défini » comme dans « Paris est la capitale de la France ».

D'après la *GMF* (2018 : 423-424), nous pouvons schématiser la phrase attributive sous la forme : « N – être – X ». Dans ce cas, le sujet N renvoie au référent à caractériser et l'attribut X, introduit par le verbe « être », concerne une caractéristique de N. Ce verbe²² fonctionne alors comme une copule qui signale l'existence d'une relation prédicative entre les deux éléments N et X. La *GMF* (2018 : 426-427) distingue ainsi trois types différents de caractérisation selon la catégorie notionnelle dénotée par l'attribut. La caractérisation du sujet est alors présentée différemment lorsque l'attribut X est :

- un nom employé sans déterminant. Il indique une propriété inhérente, l'état ou le statut du sujet (personne) : « Elle est étudiante en linguistique. »

²¹ Cf. le chapitre 2 de Schnedecker (2006).

²² Les SN attributifs peuvent également être introduits par d'autres verbes comme « devenir », « rester », « sembler », etc. Le choix peut être varié selon l'intention aspectuelle ou modale de l'expression.

- un SN précédé de l'article indéfini. Il identifie le sujet comme une occurrence du type dénoté par le nom : « Ce livre est un classique de la linguistique. »

- un SN précédé de l'article défini. Il entretient une relation d'équivalence référentielle avec son sujet : « Paris est la capitale de la France. »

Les deux premiers types évoqués plus haut ne concernent pas la coréférence car leur SN attributif décrit un aspect de leur SN sujet. Dans ce sens, ils peuvent davantage être considérés comme des cas de « near-identity » de Recasens *et al.* (2010) précédemment mentionnés et ne seront donc pas annotés comme vu à la section 2.1 *supra*. Un simple test de réversibilité nous aidera à discerner les coréférences parmi les divers types de constructions nominales attributives : autrement dit, seuls les cas où les N et X dans « N – être – X » peuvent être permutés avec le schéma « X – être – N » en satisfaisant la condition d'exclusivité sont des coréférences. Regardons l'exemple suivant :

(20) [L'absence de ce nez]_a est [une chose horrible à voir]_b. — (« Fantôme », paragraphe #3)

(20') (version suite à permutation) ? Une chose horrible à voir est l'absence de ce nez.

Dans cet exemple, le SN indéfini « une chose horrible à voir » de l'exemple (20) joue un rôle identifiant et définitoire pour son sujet « l'absence de ce nez ». Toutefois, si on permute les deux éléments comme dans l'énoncé (20'), leur relation n'est plus la même parce qu'ils ne sont pas exclusifs l'un de l'autre. Autrement dit, il peut y avoir plusieurs candidats pour « une chose horrible à voir ».

En revanche, le troisième cas présenté dans la *GMF*, « Paris est la capitale de la France », semble convenir à notre propos puisque les deux SN sont coréférentiels et permutable. Cependant, ils se basent sur une relation asymétrique, à savoir que le SN attribut (*la capitale de la France*) a une force référentielle inférieure au SN sujet (*Paris*) (*ibid.* : 427). Cela étant dit, il n'en demeure pas moins qu'ils établissent un lien coréférentiel et que la permutation des deux SN (*La capitale de la France est Paris*) ne change ni leur relation d'équivalence référentielle ni la valeur sémantique de la phrase comme l'illustre l'exemple ci-dessous extrait de notre corpus :

(21) Mademoiselle, [je]_a suis [le petit enfant qui est allé ramasser votre écharpe dans la mer]_a. — (*Fantôme*, paragraphe #154)

(21') Mademoiselle, [le petit enfant qui est allé ramasser votre écharpe dans la mer]_a est [moi]_a.

Ainsi, dans les énoncés (21) et (21'), la relation entre les deux éléments est exclusive et coréférentielle. Nous considérerons donc des constructions attributives de ce type comme annotables.

2.2.6 Les appositions nominales

Suivant Neveu (2000 : 3), l'apposition peut être comptée comme une des catégories fonctionnelles les plus récalcitrantes de la grammaire traditionnelle. Dans le manuel DEMOCRAT, la notion d'*apposition* est simplifiée pour une raison de commodité : « il y a apposition dès lors qu'il y a virgule ». N'étant pas d'accord avec ce choix qui ne se trouve pas être applicable au coréen, nous présenterons, par conséquent, les descriptions définitoires qui concernent les appositions nominales, appuyant ainsi notre position.

2.2.6.1 Constructions appositives intégrées

Rioul (1983) distingue deux types de construction appositive : intégrée et détachée. Le premier concerne des SN tels que « des années 1950 », « la ville de Paris », « ce rôle de Juliette », *etc.*, et le deuxième, les constructions qui sont accompagnées d'une virgule et qui ont la possibilité de déplacement parmi les SN. Pour les appositions intégrées qui se basent sur une équivalence référentielle, comme « ce rôle de Juliette », où « ce rôle » et « Juliette » renvoient à un même référent, nous avons décidé d'annoter une seule fois le SN entier, afin d'éviter une redondance sémantique.

En coréen, la construction appositive intégrée peut se présenter sous les formes suivantes : « N1 + -ui + N2 », « N1 + N2 » et « relative + N ». Lorsque ces SN sont employés de manière appositive, nous annotons seulement le SN entier, comme nous faisons pour le français.

2.2.6.2 Constructions appositives détachées

En ce qui concerne la notion de construction appositive détachée (ou apposition, tout simplement), nous suivons Combettes (1998a : 24-28), qui distingue l'apposition de la construction détachée, par le fait que la première recouvre les SN pourvus d'un déterminant. En suivant cette définition, nous avons décidé de ne pas annoter la construction détachée dont le SN ne possède pas de déterminant. L'absence de détermination entraînant le manque

d'autonomie référentielle, dans l'énoncé (22), l'expression « chef machiniste » n'est ainsi pas annotable. Autrement dit, les expressions de ce type ne fournissent que l'information sur la catégorie d'individus qu'elles concernent et donc la référence n'est pas suffisamment précise.

(22) La vérité est que l'imagination du squelette était née de la description qu'avait faite du fantôme, **Joseph Buquet**, *chef machiniste*, qui, lui, l'avait réellement vu. — (*Fantôme*, paragraphe #18)

En revanche, en ce qui concerne la nature référentielle d'une construction appositive, chaque constituant peut fonctionner de manière autonome. Ainsi, dans l'exemple (23) ci-dessous, « MM.Debienne et Poligny » et « les directeurs démissionnaires de l'Opéra » désignent le même référent.

(23) ...[MM.Debienne et Poligny]_a, **les directeurs démissionnaires de l'Opéra**, donnaient [leur]_a dernière soirée de gala... — (*Fantôme*, paragraphe #3)

Nous pouvons diviser cette phrase en deux énoncés comme dans (23') afin de démontrer que les deux SN sont respectivement indépendants du point de vue référentiel :

(23') a. **MM.Debienne et Poligny** donnaient leur dernière soirée de gala...

b. **Les directeurs démissionnaires de l'Opéra** donnaient leur dernière soirée de gala...

Cependant, dans la composition appositive où les deux sujets de (23'a) et (23'b) se rejoignent dans une même construction que dans (23), la hiérarchisation de l'information entre les propositions appositives est faite par l'énonciateur ce qui est généralement présenté par l'ordre d'apparition des expressions. D'après Combettes (1998a : 25, 1998b :113), le groupe placé en premier est le sujet et celui qui suit est le constituant détaché, ce dernier ayant une valeur adjectivale qui fournit une description sur le sujet. Tenant compte de ce fait, dans les constructions appositives qui ont souvent pour but de donner les informations supplémentaires sur un sujet, seule la première expression de la construction appositive sera annotée dans notre travail.

En coréen, contrairement à ce qui se passe en français, les constructions détachées ou appositives ne sont pas introduites par une virgule. Ainsi, en coréen une apposition est une phrase constituée de deux propositions indépendantes comprises l'une dans l'autre et liées entre elles par certains types de noms susceptibles de définir le caractère du contenu qui va

suivre²³ : l'aveu, la commande, la rumeur, la nouvelle, le bruit, la réclamation, la demande, la question, l'information, la résolution, la demande, l'objectif, le fait, le courage, l'expérience, etc.²⁴ Regardons l'exemple (24) ci-dessous qui illustre notre propos :

(24) 여학생이 동혁이를 찾아왔다는 소문이
 yeohagsaeng+i DongHyuki+reul chaja-watda+neun somun+i
 Etudiante+cas-suj. DongHyuk+cas-obj. chercher-venir+term.GH nouvelle+cas-suj.
 하룻동안에 동네에 파다하게 돌았다.
 harus-dongan+e dongne+e padahage dol+at+da.
 Une journée-en+cas-adv. village+cas-adv. partout se répandre+passé+term.

La nouvelle selon laquelle **une étudiante est venue chercher DongHyuk** s'est répandue partout dans le village en une journée.

Comme le montre cet exemple, les deux propositions sont indépendantes et donc séparables : « une étudiante est venue chercher DH » et « la nouvelle s'est répandue partout dans le village en une journée ». À la fin de la première phrase, la terminaison « -neun » surlignée explicite la relation appositive entre cette proposition et le nom « somun (la nouvelle) ». Cette terminaison qui lie la première proposition à la deuxième transforme en même temps la première en forme épithète (relative) qualifiant ainsi le nom « somun ». Cette sorte de construction est formellement identique aux SN ayant une construction relative comme épithète soit « relative + N ». De ce fait nous ne distinguons pas les deux cas. L'ensemble de la première proposition et le nom appositif sera annoté : [Nouvelle_YSchercheDH 여학생이 동혁이를 찾아왔다는 소문이 (la nouvelle selon laquelle une étudiante est venue chercher DongHyuk)].

2.2.7 Les anaphores résomptives

Parfois, une expression référentielle comme le pronom, le SN démonstratif ou autre peut désigner ce qui vient d'être ou va être mentionné.

²³ Les constructions de ce type sont explicitées par le biais de la notion de *sous-spécification* en français. Les noms sous-spécifiés balisent le discours (cf. articles dans la revue *Corela*, HS-34, paru en 2021).

²⁴ A l'aide d'une terminaison qui s'ajoute à la fin de la première proposition (le contenu ou la précision du nom concerné), les deux propositions sont liées.

C'est le cas de l'anaphore résomptive où une expression référentielle résume ou « encapsule » une partie du texte antérieur en créant un nouveau référent discursif (Lundquist et al., 2012 : 4). Regardons l'exemple suivant :

(25) Le fantôme leur était apparu sous les espèces d'un monsieur en habit noir qui s'était dressé tout à coup devant elles, dans le couloir, sans qu'on pût savoir d'où il venait. **Son apparition** avait été si subite qu'on eût pu croire qu'il sortait de la muraille — (*Fantôme*, paragraphe #12)

Le SN « son apparition » qui est ici l'anaphorique, renvoie à la phrase précédente qui décrit la manière dont le fantôme apparaît ; autrement dit, cette expression référentielle résume ce qui vient d'être dit dans le texte comme dans cet exemple :

(26) Enfin, Jammes s'étant jetée avec les marques d'un sincère effroi jusque dans le coin le plus reculé de la muraille, murmura [**ce seul mot**]_a : « Écoutez ! » — (*Fantôme*, paragraphe #32)

Ici, le SN démonstratif « ce seul mot » désigne l'énonciation du personnage qui suit : « Écoutez ! ». Il est parfois compliqué de borner les unités linguistiques de ce type car elles peuvent correspondre à un segment plus long et s'étendre à plusieurs phrases, un paragraphe, voire quelques pages selon les cas. Par souci de commodité, et comme proposé par le manuel DEMOCRAT, nous excluons de l'annotation l'unité linguistique qui est l'objet de la référence. Cependant, il n'en demeure pas moins qu'il existe un lien coréférentiel entre l'expression référentielle et son référent. Afin de marquer cette coréférence (seulement dans les cas où il y a au moins trois mentions coréférentielles), nous indiquons sur notre page d'annotation de SACR, l'étiquette, avec le numéro de la phrase ou celui du paragraphe, à laquelle l'objet de la référence appartient ; dans ces cas, nous attribuons le code spécial « _R » à la fin de l'étiquette pour indiquer qu'il s'agit de l'anaphore résomptive.

2.3 Les pronoms

Les divers pronoms qui peuvent renvoyer à des noms, des SN, et parfois des propositions sont généralement annotés. Le manuel DEMOCRAT décide d'annoter les pronoms personnels, démonstratifs, relatifs, adverbiaux et les pronoms indéfinis ; il exclut les pronoms réfléchis clitiques, ceux qui figurent dans des expressions figées, les pronoms impersonnels et négatifs de l'annotation. Nous suivons, dans la plupart des cas, ces critères. Cependant, vu les disparités inter-langues, il y a des exceptions à considérer (il faudra notamment noter

que certains usages des pronoms en français n'existent pas en coréen) et certains cas méritent réflexion. Ce sont les cas spécifiques dont nous allons traiter dans les pages qui suivent.

2.3.1 Pronoms personnels déictiques employés en différents niveaux discursifs

Les pronoms personnels déictiques, ainsi que *je*, *vous*, *nous* ou parfois *on*²⁵, sont employés dans plusieurs situations. Schnedecker (1997 : 12) catégorise trois niveaux discursifs : le niveau discursif de l'histoire présentée par le narrateur, le niveau discursif des commentaires méta-narratifs de l'auteur, et enfin celui des propos tenus par les personnages. Dans certains genres de textes, surtout dans les romans, cette différenciation est indispensable :

(27) Ce fer à cheval comme toute cette histoire, du reste, – hélas ! – [Narrateur **je**] ne l'ai point inventé, et [Narrateur-Lecteur **l'on**] peut encore **aujourd'hui** le voir sur la table du vestibule, devant la loge du concierge, quand [Narrateur-Lecteur **on**] entre dans l'Opéra par la cour de l'administration.

Voilà qui donne assez rapidement un aperçu de l'état d'âme de ces demoiselles, le soir où [Narrateur-Lecteur **nous**] pénétrons avec elles dans la loge de la Sorelli.

« C'est le fantôme ! » s'était donc écriée la petite Jammes. — (*Fantôme*, paragraphe #29, #30 et #31)

Dans l'exemple (27), les pronoms personnels déictiques, *je*, *on* et *nous*, sont employés au niveau discursif des commentaires méta-narratifs, désignant ainsi respectivement « le narrateur » et « le narrateur et le lecteur dans l'ensemble ». La marque temporelle « aujourd'hui » indique, de plus, une certaine distance temporelle entre l'événement du récit, représenté par l'expression « le soir où nous pénétrons avec elles dans la loge de la Sorelli », et le moment de la narration.

Comme notre corpus est constitué de romans rédigés avec le point de vue omniscient du narrateur, souvent son intervention s'insère au sein du déroulement événementiel. Lorsque le narrateur intervient, le temps du récit s'arrête et le temps propre de la narration commence,

²⁵ Il n'y a pas de pronom équivalent de « on » en coréen. La description sur le pronom « on » (comprenant la section 2.3.4 *infra*) est donc valable seulement pour le français.

provoquant chez les lecteurs un déplacement cognitif du « centre déictique »²⁶. Les lecteurs essaient, en effet, de trouver un repère spatio-temporellement correspondant à l'espace-temps de l'énonciation en cours. En lisant l'intervention du narrateur, ils cherchent donc à nouveau le centre déictique, ce repère nouvellement proposé. Tenant compte de ce fait, nous discernons donc les déictiques énoncés par le narrateur et ceux énoncés par les personnages du récit.

2.3.2 Pronom réfléchi *se* du français

Les pronoms compléments qui sont coréférentiels aux sujets dans des constructions pronominales réfléchies du verbe, peuvent être interprétés de deux manières différentes, l'une étant en effet le fait de pronoms réfléchis, l'autre, quant à elle, le fait des réciproques, comme le montrent les deux exemples qui suivent :

(28) Le comte, derrière Raoul, **se disait** : « Le coquin connaît le chemin ! »... — (*Fantôme*, paragraphe #141)

(29) Les deux directeurs, ou plutôt ex-directeurs, **se levèrent** aussitôt, en fixant étrangement leur interlocuteur. Ils étaient agités plus que de raison, c'est-à-dire plus qu'on a raison de l'être par l'annonce de la pendaison d'un chef machiniste. Ils **se regardèrent** tous deux. Ils étaient devenus plus pâles que la nappe. — (*Fantôme*, paragraphe #209)

Les pronoms réfléchis des verbes « dire » de (28) et « lever » de (29) renvoient respectivement aux agents du procès, « le comte » et « les deux directeurs ». L'interprétation est donc réfléchie. En revanche, *se* du verbe « regarder » dans (29) réfère à « ils », et donc à un ensemble avec un sens réciproque.

Par ailleurs, en coréen, ces pronoms n'ont pas de caractère clitique au verbe, mais fonctionnent plutôt comme d'autres pronoms personnels ou déterminants possessifs dans la phrase. Dans ces cas, la réciprocité s'exprime par l'adverbe « *Seolo* », non par le pronom. En conséquence, afin d'équilibrer la comparaison, nous décidons de ne pas annoter le pronom réfléchi du verbe en français.

²⁶ D'après Kleiber (2003 : 39-40), le terme « centre déictique » est un concept emprunté au psychologue et linguiste allemand Karl Bühler (il a utilisé le terme *Origo*) qui veut désigner un point de départ où le locuteur et l'interlocuteur sont posés pour appréhender les référents des expressions déictiques utilisées.

2.3.3 Pronom *il* impersonnel du français

Le pronom *il* employé avec les verbes impersonnels ou dans les locutions verbales impersonnelles²⁷ ne sera pas annoté. En effet, celui-ci est utilisé avec certains verbes dénotant un phénomène météorologique comme *pleuvoir, geler, neiger, bruiner, tonner, et venter*, ou dans certaines locutions verbales impersonnelles telles que *il faut, il s'agit de, il y a, etc.* comme dans l'exemple suivant :

(30) Il semblait, en effet, à tout le monde qu'un frôlement se faisait entendre derrière la porte. — (*Fantôme*, paragraphe #33)

Ce type d'emploi du pronom *il* est exclu de l'annotation étant donné qu'il n'a « aucune des propriétés interprétatives des sujets ordinaires et en particulier du pronom personnel *il(s), elle(s)* dans ses emplois anaphoriques et déictiques » (Riegel *et al.*, 2018 : 745).

2.3.4 Pronom *on* du français

Le pronom *on*, quant à lui, est un élément linguistique complexe à analyser car son interprétation peut varier selon les cas. Landragin & Tanguy (2014) expliquent son caractère complexe comme suit :

Il peut référer aussi bien à une personne précise qu'à un générique, à un collectif clairement identifiable qu'à un groupe de personnes à l'étendue et aux limites vagues ; il peut inclure ou exclure le locuteur, renvoyer aussi bien aux interlocuteurs qu'à de tierces personnes. (Landragin & Tanguy, 2014 : 99)

Ainsi, *on* est parfois considéré comme pronom personnel quand il est employé pour désigner « je, tu, il, nous, vous, ils », cependant, il peut aussi jouer un rôle de pronom indéfini quand il est utilisé pour épisodiquement faire référence à « quelqu'un, des gens » ou bien de façon générique à « chacun, les gens ».

2.3.4.1 *On* et ses diverses interprétations

Parfois, l'interprétation de « on » varie selon le contexte. En effet, le référent auquel il renvoie peut être soit une personne précise (exemple 31 ci-dessous), soit un générique (32),

²⁷ En coréen, le même phénomène réalisé par les verbes impersonnels ou les locutions verbales impersonnelles ne se présente pas.

un collectif clairement identifiable (33) ou bien encore un groupe de personnes à l'étendue et aux limites vagues (34) :

(31) « Ne trouvez-vous point, docteur, que ces messieurs devraient “dégager” un peu la loge ? demanda Raoul avec une incroyable audace. [_{Daaé} **On**] ne peut plus respirer ici. — (*Fantôme*, paragraphe # 144)

(32) A Paris, [_{LesGensAParis} **on**] est toujours au bal masqué... — (*Fantôme*, paragraphe # 193)

(33) À ce moment, [_{DanseusesLogeS} **on**] entendit des pas puissants et pressés dans le couloir et une voix essoufflée qui criait : — (*Fantôme*, paragraphe #84)

(34) Sa poitrine lui faisait mal, comme si [_{RF1} **on**] la lui avait ouverte pour lui prendre le cœur — (*Fantôme*, paragraphe # 139)

Le pronom « on » de l'exemple (31) est employé dans le contexte où *Christine Daaé*, s'est évanouie et est transportée dans sa loge. En fonction de ce contexte, il est raisonnable de déduire qu'ici « on » renvoie à une personne précise, *Daaé*. En revanche, selon les cas, « on » désigne aussi un référent générique comme dans (32). Ensuite, le pronom « on » réfère également à un collectif identifiable comme illustré dans l'exemple (33) où la situation contextuelle renvoie à une douzaine de danseuses du corps de ballet qui parlent entre elles du fantôme dans une loge où elles sont réunies. Cependant, dans l'exemple (34), le pronom « on » désigne quelqu'un ou des personnes indéterminées.

À part cette distinction, « on » inclut ou exclut le locuteur et l'interlocuteur. Nous reprenons l'exemple (27) *supra* où, comme nous l'avons déjà remarqué, « on » désigne chaque fois l'ensemble des narrateurs et lecteurs.

(27) Ce fer à cheval comme toute cette histoire, du reste, – hélas ! – [_{Narrateur} **je**] ne l'ai point inventé, et [_{Narrateur-Lecteur} **l'on**] peut encore aujourd'hui le voir sur la table du vestibule, devant la loge du concierge, quand [_{Narrateur-Lecteur} **on**] entre dans l'Opéra par la cour de l'administration...

Dans de nombreux cas, le contexte comprenant divers éléments du texte tels que le prédicat, l'adverbe cadratif et les marqueurs morphologiques ou sémantiques peut contraindre ou suggérer une certaine interprétation de *on*. Dans (27), cette référence est saisie à partir du contexte, par exemple par un repère temporel comme « aujourd'hui » et par des prédicats employés au présent.

Landragin & Tanguy (2014 : 105) soulignent également l'importance du contexte en empruntant une partie de la thèse de Gjesdal (2008 : 12) :

[...] Il semble plutôt que le sens de ON soit attribué à partir d'une interaction entre des éléments inhérents et des facteurs contextuels. (Gjesdal, 2008 : 12)

Pourtant, dans des cas où l'ambiguïté ne peut être levée malgré les indices contextuels fournis, nous attribuons des étiquettes du type RF1, RF2...²⁸ (forme abrégée de 'RéfèrentFlou1') comme dans l'exemple (34).

2.3.4.2 La forme l'on

La forme « l'on », constituée de « l'apostrophe » et du pronom *on* a pour rôle de faciliter la prononciation dans certaines constructions. Dans notre travail d'annotation, nous la considérons comme un seul bloc comme suit :

(35) ...nous allons aux nouvelles d'un pauvre homme que [_{Machinistes} l'on] a trouvé pendu.
— (*Fantôme*, paragraphe #107)

2.3.5 Pronoms relatifs

Le coréen ne possède pas de pronom relatif. Comme nous venons de le mentionner *supra* (voir la section 4.4.1.4 du chapitre 4), la proposition relative antéposée dans le SN est liée au nom par une terminaison « -ㄴ (-n) ». Cette partie sur le critère d'annotation des pronoms relatifs ne sera donc valable que pour le français.

²⁸ Les chaînes de ce type seront malgré tout assez rares car si le pronom « on » est flou, cela veut dire qu'il est difficile de décider la limite de son référent. Le pronom « on » employé respectivement dans les deux exemples suivants peut être interprété comme renvoyant « aux gens en général » de limite vague :

(ex1) **On** dit communément d'elle que c'est 'une belle créature'. — (*Fantôme*, paragraphe #43).

(ex2) Les petites danseuses se regardèrent. Si le fantôme venait dans la loge, **on** devait le voir... — (*Fantôme*, paragraphe #80).

Cependant, il est difficile de dire si ce « on » employé dans (ex1) est coréférentiel avec celui dans (ex2). Même s'ils désignent tous « les gens en général », leurs référents réels peuvent varier selon le contexte dans lequel ils sont apparus : d'après son contexte, il est fort probable que dans (ex2), « on » réfère aux gens comprenant « les petites danseuses » tandis que le contexte (ex1) ne revendique pas cette sorte d'inclusion obligatoire d'un certain groupe. Il est donc important d'examiner chaque emploi et son référent en fonction du contexte.

Le traitement des pronoms relatifs du français, quant à lui, est fait en suivant le manuel DEMOCRAT.

Tout d'abord, intéressons-nous au cas où les pronoms relatifs introduisent une construction relative en établissant un lien coréférentiel avec le référent qui les précède (Riegel *et al.*, 2018 : 795). Les exemples suivants mettent en évidence les pronoms relatifs qui relient les deux propositions :

(36) Au premier palier, la Sorelli se heurta [_{Chagny} **au comte de Chagny**] [_{Chagny} **qui**] montait.
— (*Fantôme*, paragraphe #105)

(37) Mais celui-là avait [_{Spécialité} **une spécialité**] [_{Spécialité} **que**] n'ont point tous les habits noirs. Il habillait un squelette. — (*Fantôme*, paragraphe #15)

(38) Eh bien, aussitôt que la porte s'est ouverte devant le Persan, Gabriel ne fit qu'un bond [_{Fauteuil} **du fauteuil**] [_{Fauteuil} **où**] il était assis jusqu'à la serrure de l'armoire, pour toucher du fer ! — (*Fantôme*, paragraphe #55)

Considérant le fait que ces pronoms sont coréférents à leur antécédent, les relatives entières ne sont alors pas comprises dans les bornes des mentions comme nous le voyons dans les exemples (36), (37) et (38) : seuls le pronom relatif et son antécédent sont annotés.

Ce qui nous semble complexe au niveau de la délimitation des mentions, ce sont les cas où les pronoms relatifs n'ont pas d'antécédent mais sont encore suivis d'une relative. Le manuel DEMOCRAT classe ce deuxième cas sous le nom de « pronoms sans antécédent » et annote l'ensemble comme une seule mention. Regardons les exemples suivants :

(39) [**Qui n'a pas entendu Christine chanter le trio final de Faust**]_a ne connaît pas Faust. — (*Fantôme*, paragraphe #117)

(40) Et voici ce que Joseph Buquet a dit du fantôme à [**qui voulait l'entendre**]_b — (*Fantôme*, paragraphe #19)

Ici, étant donné qu'il n'y a pas d'antécédent, toute la proposition relative est annotée (seulement si elle arrive à construire une chaîne) afin de bien identifier le référent ; le pronom relatif tout seul, lui, n'est pas annoté à part.

2.3.6 Pronoms démonstratifs et constructions pseudo-clivées

De leur côté, les pronoms démonstratifs font l'objet d'une annotation. En effet, il est

important de remarquer que la façon de définir les pronoms démonstratifs diffère selon la langue : contrairement au français, en coréen, ils sont classés en deux sous-catégories, les démonstratifs renvoyant aux choses et aux lieux²⁹.

Le principe d'annotation est le même : nous les annotons lorsqu'ils établissent un lien coréférentiel avec d'autres entités du texte. Dans les sections suivantes, nous présentons rapidement les deux cas de pronom démonstratif du français qui méritent discussion.

2.3.6.1 Forme simple et forme renforcée

Voici deux formes différentes sous lesquelles ils peuvent apparaître dans le texte :

(17) Et comme elle était la première à croire [_{Fantôme_G} **aux fantômes en général**] et à [_{Fantôme_Opéra} [_{Fantôme} **celui**] de l'Opéra] en particulier, elle voulut tout de suite être renseignée. [...] — (*Fantôme*, paragraphe #7)

(41) MM. Debienne et Poligny, placés au milieu de la table, n'avaient pas encore aperçu [_{Fantôme} **l'homme à la tête de mort**], quand [_{Fantôme} **celui-ci**] se mit tout à coup à parler. — (*Fantôme*, paragraphe #205)

« Celui » dans l'exemple (17) *supra* est classé comme une forme simple du pronom démonstratif (comme *celle* et *ce* qui appartiennent à la même catégorie). La *GMF* (2018 : 376) souligne qu'il « reprend le contenu lexical et le genre d'un nom antécédent » en ayant la possibilité de modifier le nombre et les déterminations en conformité avec son but référentiel. Pour cela, le pronom démonstratif d'une forme simple est toujours suivi d'un complément.

Ensuite « celui-ci » de (41), est la forme renforcée de pronom démonstratif : la forme simple « celui » est suivi de particules adverbiales *-ci* (ou *-là* selon les cas). Il peut être employé de manière autonome, c'est-à-dire que contrairement à la forme simple, il n'a pas besoin d'être complété. D'après la *Grammaire du français* (Denis et Sancier-Chateau, 1994 : 188), les pronoms démonstratifs de cette forme peuvent fonctionner de deux manières

²⁹ Dans chaque sous-catégorie, trois démonstratifs sont distingués selon la distance entre l'objet et le locuteur (ou son interlocuteur).

	Proche du locuteur	Loin du locuteur	Loin du locuteur mais proche de l'interlocuteur
Chose	igeos (celui-ci)	jeogeos (celui-là)	geugeos (celui-là)
Lieu	yeogi (ici)	jeogi (là)	geogi (là)

différentes : en tant que représentants qui reprennent un élément (ou bien son contenu notionnel) déjà mentionné dans le contexte comme c'est le cas dans l'exemple (41) et comme nominaux qui portent une valeur déictique et trouvent le référent dans la situation de discours comme « cela » dans l'exemple (42) :

(42) « Comment pouvez-vous me dire **cela** ? Moi qui ne chante que pour vous ! » —
(*Fantôme*, paragraphe #167)

Le pronom « cela » ici renvoie à ce que l'interlocuteur a dit dans la situation précédente. Autrement dit, il anaphorise un segment textuel plus large de manière résomptive (voir la section 2.2.7 *supra*).

2.3.6.2 C'est et les constructions pseudo-clivées

Les pronoms démonstratifs neutres (« ce³⁰ » pour la forme simple, « ceci », « cela », et « ça » pour la forme renforcée) doivent être annotés de manière générale comme dans l'exemple ci-dessous :

(43) [LieutenantDePompiers **Un lieutenant de pompiers**], [LieutenantDePompiers **c'**est brave !
[LieutenantDePompiers **Ça**] ne craint rien, [LieutenantDePompiers **ça**] ne craint surtout pas le feu ! —
(*Fantôme*, paragraphe #24)

Le manuel DEMOCRAT (2017 : 13-14) annoté « c' » et le N dans la construction « c'est (DET) N » car il est coréférent au N alors qu'il ne l'est pas dans « c'est ADJ ». Cependant, dans cette recherche, nous annotons « c' » des constructions de « c'est ADJ » sauf dans le cas où la locution « c'est » est employée comme un simple présentatif sans référence précise. Ainsi, par exemple, dans la phrase « Un lieutenant de pompiers, c'est brave » tirée de l'exemple ci-dessus, où le pronom démonstratif renvoie à « un lieutenant de pompiers », il sera donc annoté.

Par ailleurs, « c'est » peut apparaître dans des constructions clivées et pseudo-clivées. D'après la définition proposée par la *GMF*, ce sont des constructions dont « un constituant est extrait de la phrase et placé au début de celle-ci, encadré par *c'est* et par le pronom relatif *qui* ou *que* » (Riegel *et al.*, 2018 : 725). Elles ont pour objectif de mettre en relief tel ou tel élément

³⁰ Le pronom démonstratif « ce » s'emploie comme sujet clitique du verbe « être ».

de la phrase³¹.

(44) C'était [_{Persan} **le Persan**] qui entrait. — (*Fantôme*, paragraphe #53)

(45) Et c'était de **cela** qu'il voulait parler à la Daaé, sans retard. — (*Fantôme*, paragraphe #165)

L'élément extrait, quel que soit son rôle dans la phrase, se situe entre « c'est » et « qui » ou « que ». Par exemple, la phrase (44) est la construction clivée de « Le Persan entrait » et la (45) correspond à celle de « Il voulait parler de cela à la Daaé, sans retard ». Vu ces phrases reformulées, nous ne considérons pas qu'il y ait la coréférence entre « c'est » et le pronom relatif : il s'agit d'un dispositif morpho-syntaxique du focus³².

Il nous faut ici ajouter un exemple qui paraît semblable mais qui, pourtant, doit être catégorisé différemment : la phrase pseudo-clivée :

(46) [**Tout ce qu'on a raconté sur sa tête de mort et sur sa tête de feu**], [c]'est [**des blagues**] ! — (*Fantôme*, paragraphe #81)

Les constructions de ce type, catégorisées comme « thematic equatives » (Halliday : 2004, cité par Rouquier : 2018), ont une structure détachée « A, c'est B ». Il s'agit d'une combinaison de « l'extraction » et « du détachement ». Nous considérons dans ces cas qu'il y a une coréférence entre « A », « c' », et « B ».

2.4 Les verbes

En général, les verbes ne sont pas considérés comme des éléments annotables. Cependant, nous annotons les verbes dont le sujet est le pronom zéro.

2.4.1 Coordination des verbes

La coordination de plusieurs propositions ayant le même sujet concerne le pronom zéro comme le décrit l'exemple (47) :

(47) [_{Ex-directeurs} **MM. Debienne et Poligny**] l'avaient embrassée, [_{Ex-directeurs} **remerciée**] et

³¹ Ce point sur les constructions clivées est valable uniquement pour le texte français : la construction clivée n'existe pas en coréen.

³² Cf. Rouquier (2018 : 13-14).

[Ex-directeurs **s'étaient**] sauvés aussi rapides que le fantôme lui-même. — (*Fantôme*, paragraphe #197)

Dans cette phrase, les trois actions différentes sont coordonnées dans l'ordre chronologique. Le sujet n'y est exprimé qu'une fois au début de la phrase parce que ces actions sont toutes réalisées par les mêmes personnages, « MM. Debienne et Poligny ». Dans ce type de cas où le pronom zéro est employé, nous annotons le verbe.

2.4.2 Destinataire de l'expression impérative

Selon la *GMF* (2018 : 575), l'impératif est « un mode personnel et non temporel qui est restreint en personnes », c'est-à-dire qu'il ne concerne que la première personne du pluriel et les deuxièmes personnes du singulier et du pluriel car il est réservé à l'expression d'une valeur directive telle que l'ordre, l'exhortation, le conseil, etc. En règle générale, dans les expressions impératives, le sujet est structurellement absent : la désinence verbale spécifie la personne visée par le locuteur. D'après les auteurs de *Grammaire du français* (1994 : 221), c'est en raison de la présence du destinataire dans la situation d'énonciation que sa formulation explicite devient inutile.

Les expressions impératives sont fréquemment présentes dans le discours du personnage de texte littéraire comme dans l'exemple suivant :

(48) – Oh ! oui ! Oui ! certainement, il y a quelqu'un derrière la porte ! » répéta ce petit pruneau sec de Meg Giry, qui retint héroïquement la Sorelli par sa jupe de gaze... « Surtout, [_{LaSorelli} **n'ouvrez pas !**] Mon Dieu, [_{LaSorelli} **n'ouvrez pas !**]» — (*Fantôme*, paragraphe #36)

D'après le contexte, l'impératif en gras dans l'exemple est une prière du personnage « Meg Giry » dont le destinataire est « la Sorelli ».

Nous reconnaissons que le traitement de l'impératif est un sujet polémique dans l'annotation des CR. Le manuel DEMOCRAT propose d'annoter les expressions de ce type parce qu'il les considère comme le cas du pronom zéro. Malgré le fait que cette idée soit souvent contestée, nous choisissons d'annoter le verbe à l'impératif dans cette thèse. Ce choix se base sur le point de comparaison avec le coréen où l'on considère, dans la majorité des cas, que la construction impérative contient le pronom zéro (Nam Kisim *et al.*, 2019 : 267-270).

2.5 Autres cas

Nous mentionnerons rapidement les sujets qui ne sont pas traités dans les pages précédentes : l'annotation des groupes de noms coordonnés, celle des éléments partageant le même contenu lexical et le traitement des titres et des nombres.

2.5.1 Coordination et délimitation des groupes

Quand plusieurs éléments sont coordonnés, nous annotons d'abord chaque élément puis tous les éléments ensemble parce que ceux-ci peuvent être repris à la fois respectivement et en totalité comme dans le texte (49) :

(49) Songez donc que [_{Compositeurs} [Gounod **Gounod**], [_{Reyer} **Reyer**], [_{Saint-Saëns} **Saint-Saëns**], [_{Massenet} **Massenet**], [_{Guiraud} **Guiraud**], [_{Delibes} **Delibes**]], montèrent à tour de rôle au pupitre du chef d'orchestre et [_{Compositeurs} **dirigèrent**] [_{Compositeurs} **eux-mêmes**] l'exécution de [_{Compositeurs} **leurs**] œuvres. [_{Compositeurs} **Ils**] eurent, entre autres interprètes, Faure et la Krauss, et c'est ce soir-là que se révéla au Tout-Paris stupéfait et enivré cette Christine Daaé dont je veux, dans cet ouvrage, faire connaître le mystérieux destin.

[_{Gounod} **Gounod**] avait fait exécuter La marche funèbre d'une Marionnette ; [_{Reyer} **Reyer**], sa belle ouverture de Sigurd ; [_{Saint-Saëns} **Saint-Saëns**], La Danse macabre et une Rêverie orientale ; [_{Massenet} **Massenet**], une Marche hongroise inédite ; [_{Guiraud} **Guiraud**], son Carnaval ; [_{Delibes} **Delibes**], La Valse lente de Sylvia et les pizzicati de Coppélia, Mlles Krauss et Denise Bloch avaient chanté : la première, le boléro des Vêpres siciliennes ; la seconde, le brindisi de Lucrece Borgia. — (*Fantôme*, paragraphe #111 et #112)

2.5.2 Les éléments partageant le même contenu lexical

Dans le texte, on utilise souvent une astuce pour éviter la répétition quand il s'agit de deux entités différentes partageant le même contenu lexical. L'emploi du pronom démonstratif « celui » que nous avons mentionné *supra* (dans la section 2.3.6.1.) appartient également à cette catégorie.

Dans l'exemple suivant, le groupe de deux hommes « Monsieur Debiegne et Monsieur Poligny » peut être contracté en « MM.Debienne et Poligny ». Ce type d'expressions doit être

annoté ainsi :

(50) MM.Debienne et Poligny

→ [Ex-directeurs MM.[MonsieurDebienne Debienne] et [MonsieurPoligny Poligny]]

Dans les expressions identiques à celles de l'exemple (50), nous devons toutefois faire attention aux constructions où les adjectifs sont énumérés pour qualifier un seul sujet comme dans « une voix forte et curieuse ». Afin de différencier ce dernier et le (50), nous proposons les trois tests du nombre de référent, du recopiage et du « à la fois ».

(51) a. Les forêts publiques et privées

b. Une voix forte et curieuse

c. Des pas puissants et pressés

D'abord, les trois cas peuvent être distingués par le nombre de référent qu'ils désignent. Dans (51a), on parle de deux types de forêt différents alors que, dans (51b), une seule voix fait l'objet de la référence. Pourtant, dans (51c), il est difficile de dire qu'il s'agit d'un seul référent. Pour éviter cette sorte de confusion, nous pouvons recopier le nom et comparer la version recopiée avec l'originale pour voir si elles ont toujours la même valeur sémantique. La version recopiée de (51a) est « les forêts publiques et les forêts privées ». Les référents auxquels renvoient cette version sont identiques à ceux qui sont désignés par (51a). Au contraire, si nous recopions le SN (51b), l'interprétation n'est plus la même : « Une voix forte et une voix curieuse » réfère à deux voix différentes. Il en va de même pour (51c) : la version recopiée « des pas puissants et des pas pressés » désigne deux types de pas alors que la version originale veut souligner les deux qualités différentes des pas concernés.

Finalement, nous pouvons ajouter l'expression de simultanéité « à la fois » après le nom : « les forêts à la fois publiques et privées », « une voix à la fois forte et curieuse » et « des pas à la fois puissants et pressés ». Tandis que le premier est contradictoire (les forêts ne peuvent pas être publiques et privées en même temps), le deuxième et le troisième gardent la même signification de leur version originale. Par conséquent, nous annotons ces cas de manière différente comme illustré dans l'exemple ci-dessous :

(52') a. Les forêts publiques et privées

→ [Forêts_France Les forêts [ForêtPublique publiques] et [ForêtPrivée privées]]

b. Une voix forte et curieuse

→ [VoixForteEtCurieuse Une voix forte et curieuse]

c. Des pas puissants et pressés

→ [PasPuissantsEtPressés Des pas puissants et pressés]

2.5.3 Titre du texte et titre de chaque chapitre

Les titres ont pour caractéristique d'être synthétiques : ils nous fournissent l'information globale du texte avant la lecture. Le texte est considéré comme une forme d'« élaboration » du titre comme le montre l'exemple suivant qui est un extrait du roman *Fantôme* :

(52) Chapitre I : Est-ce le fantôme ?

Ce soir-là, qui était celui où MM. Debienne et Poligny, les directeurs démissionnaires de l'Opéra, donnaient leur dernière soirée de gala, à l'occasion de leur départ, la loge de la Sorelli, un des premiers sujets de la danse, était subitement envahie par une demi-douzaine de ces demoiselles du corps de ballet qui remontaient de scène après avoir « dansé » Polyeucte. Elles s'y précipitèrent dans une grande confusion, les unes faisant entendre des rires excessifs et peu naturels, et les autres des cris de terreur.

La Sorelli, qui désirait être seule un instant pour « repasser » le compliment qu'elle devait prononcer tout à l'heure au foyer devant MM. Debienne et Poligny, avait vu avec méchante humeur toute cette foule étourdie se ruer derrière elle. Elle se retourna vers ses camarades et s'inquiéta d'un aussi tumultueux émoi. Ce fut la petite Jammes, – le nez cher à Grévin, des yeux de myosotis, des joues de roses, une gorge de lis, – qui en donna la raison en trois mots, d'une voix tremblante qu'étouffait l'angoisse : C'est le fantôme ! — (*Fantôme*, paragraphe #3 et #4)

Contrairement au titre qui mentionne « le fantôme », dans la partie subséquente, l'existence du fantôme ne se révèle qu'à la fin du deuxième paragraphe. Dans les romans, considérant le fait que les titres sont souvent hors du déroulement naturel (ou bien chronologique) du récit, nous ne traitons pas les segments dans l'espace du titre comme maillons. Par conséquent, ni le titre entier ni ses parties « chapitre I », « ce » et « le fantôme » ne sont annotés dans « Chapitre I : Est-ce le fantôme ? ».

2.5.4 Les nombres

Seuls les nombres qui renvoient à une date et constituent une chaîne sont annotés (par exemple « depuis 1875 » ou « des années 1950 »). Les simples chiffres ne sont pas annotés sauf si le nombre lui-même est un thème principal du texte.

Jusqu'à présent, nous avons présenté et explicité nos critères d'annotation ainsi que les raisons nous ayant amenée à faire ces choix. Certains sont faits sur la base du manuel DEMOCRAT et d'autres en considérant leur applicabilité aux deux langues différentes. Dans la section suivante, nous rendrons compte des schémas des propriétés que nous avons adoptés afin d'aborder le résultat pertinent de l'annotation.

3 Schéma des propriétés

Le schéma des propriétés est appliqué à la troisième étape de l'annotation afin d'attribuer à chacun des maillons un ensemble de caractéristiques nécessaires pour l'analyse des données³³. En d'autres termes, le schéma montre ce que nous souhaitons savoir par l'annotation. Tenant compte du fait qu'il s'agit d'une étude contrastive, il est important de construire un schéma pertinent pour chaque langue afin de comprendre le phénomène apparu respectivement dans les deux différents systèmes langagiers et d'obtenir un résultat fiable. L'enjeu est donc de choisir les critères essentiels pour l'analyse et de les sous-catégoriser en fonction du caractère de chaque système langagier.

3.1 Critères communs

Dans cette étude, nous choisissons d'appliquer deux critères communs pour les deux langues : le critère catégoriel et le critère référentiel. Le premier nous aidera à comprendre de manière globale la distribution, la composition des CR et leurs constituants alors que le deuxième servira à une analyse qualitative susceptible de montrer les fonctions référentielles

³³ Voir la section 2.3 du chapitre 5.

de chaque maillon.

3.1.1 Critère catégoriel

Dans le chapitre 4, nous avons classé les expressions référentielles en trois catégories : les noms propres, les pronoms personnels et les SN. Cette description vise à fournir une vue globale sur les deux langues et leurs façons de référer. Cependant, afin d'analyser les textes annotés non seulement de manière quantitative mais aussi qualitative, nous sentons la nécessité de détailler de nouveau ces trois catégories en plusieurs sous-catégories afin de saisir les différentes caractéristiques de chaque type d'expression référentielle.

Considérant la discordance du classement morphologique des unités linguistiques entre le français et le coréen, il est important de choisir les concepts susceptibles de s'appliquer à toutes les expressions référentielles de deux côtés. Avant tout, nous avons donc brièvement énuméré quelques exemples concrets de types de maillons apparus dans notre texte français puis constitué une liste de termes morphologiques :

- Npr nu : *Joseph Buquet, Christine, Poligny, etc.*
- Npr modifié : *la petite Daaé, la Sorelli, un Chagny, etc.*
- SN nu : *gala, écriture d'enfant*
- SN avec déterminant (le terme juste sera *SN déterminé* pour le coréen) : *ce soir-là, ses camarades, un vrai fantôme, la mère, etc.*
- Déterminant possessif : *leur*
- Pronom personnel : *je, vous, il, les, etc.*
- Pronom démonstratif : *celui, ça, cela, etc.*
- Pronom relatif : *qui, que, dont, laquelle, etc.*
- Pronom zéro : *elle frissonna et (-) dit*
- Autres³⁴ : *y, en, tous, là, ici.*

³⁴ Cette catégorie « fourre-tout » comprend les expressions qui ne sont pas catégorisées au-dessus. En effet, les trois premiers (*en, y et tous*) sont pronominaux alors que les deux derniers (*ici et là*) sont plutôt adverbiaux. Même si la catégorie grammaticale de ces expressions n'est pas identique, ce choix nous aidera à analyser la répartition des divers types d'expressions référentielles apparus dans les textes en dehors de ceux qui sont déjà catégorisés.

En français, la détermination d'un nom dépend des éléments grammaticaux tels que l'article défini, l'article indéfini, le démonstratif et le possessif. Les noms employés avec ces éléments seront donc classés dans la catégorie « SN avec déterminant ». En revanche, comme nous venons de le voir dans les chapitres précédents, en coréen, langue sans système d'article, la notion équivalant exactement à celle de déterminant en français n'existe pas. Pourtant, la langue coréenne a bien sa propre façon, soit directe soit allusive, de déterminer les noms. En effet, certains éléments grammaticaux qui qualifient le nom peuvent y participer directement, mais, même s'il n'y a aucun élément qualificatif, d'autres constituants du contexte, comme la particule ou le prédicat, peuvent participer à la détermination du nom d'une manière allusive. Les noms déterminés directement par diverses sortes de qualificatifs seront inclus dans la catégorie « SN déterminé » que nous détaillerons par un sous-critère du schéma.

De même, lorsque nous annotons le texte coréen, nous retirons les termes morpho-syntaxiques « Déterminant possessif » et « Pronom relatif » présentés dans la liste *supra* car ce sont des notions inadaptées au coréen. En coréen, à la différence du français, il n'existe pas de catégorie grammaticale particulière exprimant la relation possessive. Ainsi, afin d'exprimer cette relation, il faut simplement utiliser un pronom ou un nom (soit Npr soit SN) de possesseur et postposer la particule « -ui ». Les expressions de ce type doivent donc être classées selon la catégorie grammaticale du nom de possesseur.

Nous retirons également « pronom relatif » de la liste puisque cette catégorie grammaticale n'existe pas non plus en coréen. Cependant, comme nous l'avons déjà précédemment montré *via* les exemples, malgré l'absence du pronom relatif, la proposition relative peut être insérée dans la phrase en tant que « embedded phrase³⁵ » et fonctionne alors comme un qualificatif. En considérant tous les aspects que nous venons d'évoquer, la liste suivante indique les termes morphologiques à appliquer lors de l'annotation du texte coréen :

- Npr nu : 동혁 (*DongHyuk*), 영신 (*YoungShin*), etc.

- Npr modifié : 이 미스터 박은 (*Ce monsieur Park*), 전 조선의 (*tout Chosun*), etc.

³⁵ La phrase de ce type est constituée de deux constructions dont chacune possède un sujet et un prédicat. La construction insérée dans la phrase originelle a un rôle équivalent ; de plus, à la place des pronoms relatifs, les terminaisons conjonctives se présentent sous diverses formes en fonction du type de prédicat afin de relier les deux constructions.

- SN nu : *사회자는 (présentateur), 대원들의 (militants), etc.*
- SN avec déterminant : *오늘 저녁은 (ce soir), 이 운동을 (ce mouvement), etc.*
- Pronom personnel : *그들의 (leur), 그는 (il), 여러분 (vous), etc.*
- Pronom démonstratif : *여기서 (ici), 이 (ce), etc.*
- Pronom zero : *여지간히 고집이 세구나 (- est vraiment têtue), etc.*
- Autres : *그리 (vers là), 이리 (vers ici), etc.*

3.1.2 Critère référentiel

Les maillons, même s'ils appartiennent à la même catégorie morpho-syntaxique, peuvent avoir une fonction référentielle différente. Nous avons déjà observé que plusieurs types d'expressions référentielles peuvent être employés afin d'effectuer une même fonction référentielle générique ou spécifique³⁶. Pour refléter ce fait dans notre travail, nous insérons le critère référentiel dans notre schéma tout en distinguant les maillons qui renvoient à des référents génériques de ceux qui renvoient aux spécifiques. Les maillons de ce dernier type seront encore subdivisés en deux sous-catégories : définis et indéfinis.

- Référence générique
- Référence spécifique – défini
 - indéfini

3.2 Critères supplémentaires

Nous définissons chaque maillon avec l'une des catégories et l'une des fonctions référentielles. Ensuite, afin d'identifier ces maillons plus méticuleusement, nous ajoutons différents sous-critères qui varient selon le caractère systémique de chaque langue.

Pour le français, les sous-critères ajoutés seront « le type de déterminant » et « le type d'expansion » : nous allons ainsi classer les SN déterminés et les Npr modifiés (dans le cas échéant) selon le type de déterminant avec lequel ils sont employés. Ensuite, pour ceux ayant ce type de structure, nous prendrons en compte la façon dont le maillon est expansé (par

³⁶ Voir la section 2.2.3 de ce chapitre.

adjectif, complément adnominal, relative et autres³⁷).

En revanche, pour le coréen, à part les classements catégorique et référentiel, nous allons distinguer les différents types d'épithètes qui qualifient et même déterminent certaines catégories grammaticales³⁸ : cela concerne les Npr modifiés, les SN déterminés³⁹, les pronoms personnels, les pronoms démonstratifs et certaines expressions référentielles classées dans la catégorie *autres*. Celles qui ne sont pas employées avec l'un de ces qualificatifs peuvent en quelque sorte être considérées comme étant déterminées en fonction du contexte par des éléments grammaticaux tels que la particule et le prédicat, et ce pour des raisons de commodité.

Ensuite, afin de faciliter la comparaison inter-langues lors de l'annotation des textes coréens, nous ajoutons un autre critère, « le caractère déterminatif », dont le choix est cette fois-ci, basé sur la réflexion syntaxico-sémantique. Ainsi, certaines expressions, même si elles ne sont pas manifestement déterminées par des éléments grammaticaux quels qu'ils soient, comme, par exemple, le SN nu, ont la possibilité d'être considérées comme ayant un caractère déterminatif. Ce choix est important lorsque l'on annote le texte coréen étant donné que cette langue contient de nombreux mots dits « sino-coréen »⁴⁰ dont chaque syllabe a une signification intrinsèque et est donc susceptible d'avoir un sens déterminatif en soi⁴¹. Les maillons du texte coréen seront alors distingués en trois sous-catégories : « démonstratif », « défini » ou « indéfini » autant qu'on le pourra, espérant ainsi combler, pour une part, le manque de critère unifié pour les deux langues.

³⁷ Ce groupe « divers » comprend la construction telle que « préposition + infinitif » comme dans « la première à croire » ou l'adverbe comme dans « aux fantômes *en général* ».

³⁸ Voir la section 2.2 du chapitre 3.

³⁹ Rappelons que les noms sont déterminés de manière soit directe soit allusive dans la langue coréenne pour combler le manque d'articles.

⁴⁰ Ce sont des mots empruntés directement au chinois ou de nouveaux mots coréens créés à partir de caractères chinois.

⁴¹ Par exemple, « geumnyeon » qui signifie « cette année » est un mot constitué de deux syllabes qui ont pour origine un caractère chinois. La première syllabe « geum » a un sens démonstratif en soi, ce qui rend le terme nu entier « geumnyeon » déterminé.

3.3 Schémas réalisés

Nous construisons deux versions du schéma pour rendre compte des deux systèmes langagiers. Cette distinction se révèle pertinente du fait que notre objectif final est de fournir le résultat le plus lisible possible de l'analyse des CR dans les deux langues. Ce premier schéma ci-dessous vise à être appliqué pour le français :

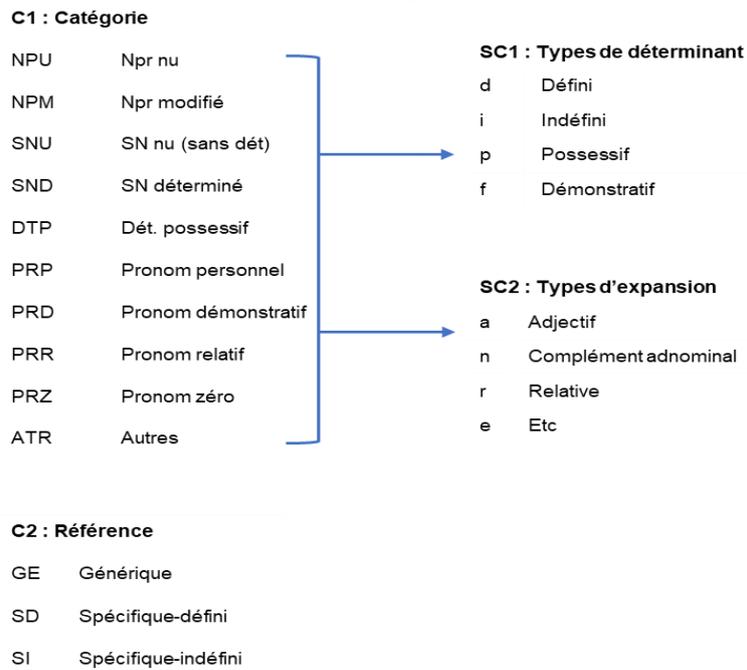


Figure 25 Schéma des propriétés pour le texte français

Ensuite, ce second schéma concerne le coréen :

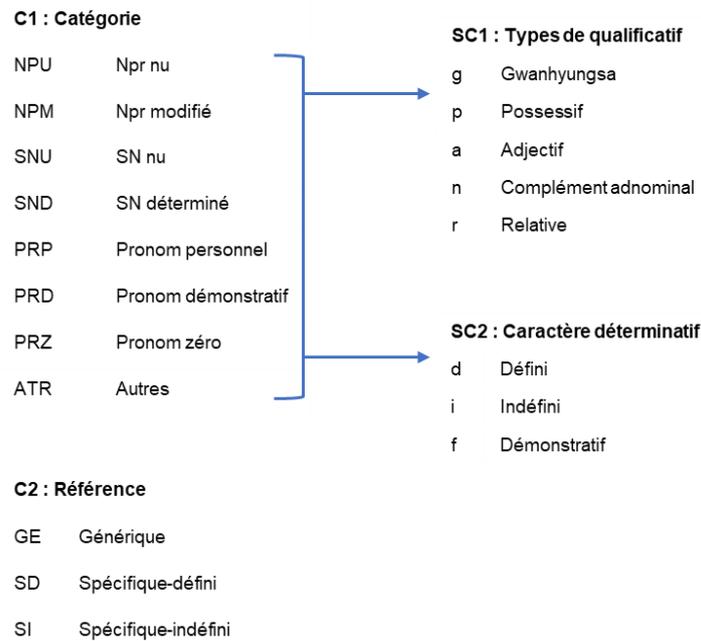


Figure 26 Schéma des propriétés pour le texte coréen

4 Résultats escomptés

Nous escomptons obtenir des données quantitatives de manière à observer à la fois la constitution des chaînes de référence, plus précisément, la proportion d'entités faisant partie de chaînes dans chaque langue, et la façon dont les diverses fonctions référentielles sont réalisées.

Nous détaillerons donc, dans le chapitre suivant, les résultats quantitatifs obtenus *via* CR Viewer. Mais avant de passer à ce chapitre, nous souhaitons présenter quelques hypothèses élaborées au fur et à mesure de l'annotation.

- H1 : le coréen n'ayant pas de genre grammatical, les pronoms personnels de troisième personne sont susceptibles d'être moins développés et moins utilisés que dans les langues indo-européennes. Par compensation, pour référer à la troisième personne, le SN déterminé ayant *gwanhyungsa* comme qualificatif, « *jisi gwanhyungsa* (démonstratif) + N » est plus couramment utilisé.
- H2 : considérant la façon dont on détermine le nom dans chaque langue, la proportion

de SN nu sera plus importante en coréen alors que le français, de son côté, favorisera théoriquement l'utilisation de SN déterminés. Cependant, le fait que, comme nous venons de le mentionner dans H1, le SN de la forme « *jisi gwanhyungsa* + N » du coréen remplace souvent le pronom personnel de la troisième personne, aurait selon nous un effet sur la proportion de SN déterminés du coréen.

- H3 : le pronom zéro sera plus souvent observé dans le texte coréen étant donné qu'on ellipse souvent le sujet ou l'objet lorsqu'on a un certain niveau de connaissance du contexte.

En analysant les textes annotés, nous examinerons si ces hypothèses sont valables ou non et présenterons la tendance de l'utilisation des expressions référentielles en français et en coréen. Au cours de ces travaux, nous espérons voir comment les CR influencent la structuration textuelle ou inversement et comment elles sont influencées par cette dernière.

Bilan

Ce chapitre, dont l'objectif premier était d'explicitier nos stratégies d'annotation, a permis de montrer la structure complètement différente des deux langues en mettant en lumière la complexité du sujet en lui-même et nos efforts pour réconcilier les particularités du français et celles du coréen.

Nous avons tout d'abord essayé de mettre en évidence nos critères d'annotation des expressions référentielles : les SN, les pronoms, les verbes et d'autres éléments textuels. Certains d'entre eux suivent les propositions du manuel DEMOCRAT et d'autres sont faits différemment soit en raison du manque d'explication satisfaisante de DEMOCRAT soit en raison de problèmes d'applicabilité au coréen. Malgré nos efforts, nous admettons que les conventions ne sont pas parfaitement satisfaisantes et qu'il reste encore des sujets à discuter. Pourtant, nous croyons que ce chapitre pourra guider les débutants dans le domaine de l'annotation des CR, notamment ceux qui veulent mener une étude inter-langues.

Ensuite, les schémas susceptibles de refléter à la fois les points communs et les disparités inter-langues sont proposés. Les critères supplémentaires que nous avons proposés permettent d'avoir plus d'exactitude dans le discernement des référents malgré l'absence d'articles dans la langue coréenne.

Enfin, suivant ces délibérations linguistiques, nous avons fait trois hypothèses que nous allons vérifier dans le prochain chapitre. Dans le chapitre qui suit qui est le dernier chapitre de ce travail, les données extraites des textes réellement annotés seront présentées et analysées.

Troisième partie :
Synthèses des analyses

Chapitre 7

Analyse des données

Introduction

L'annotation est un travail fondamental des recherches sur les CR. Comme nous l'avons souligné dans les chapitres précédents, bien que le travail soit coûteux, il s'agit d'une étape qu'il ne faut pas négliger, car cette tâche vise à fournir des données globales et à trouver des exemples précis illustrant des phénomènes linguistiques. Pour bien annoter notre corpus, plusieurs relectures ont été nécessaires. Puis, une fois le texte complètement annoté, une stratégie d'analyse adéquate a été mise en place afin de l'exploiter.

Dans ce qui suit, nous présenterons d'abord des chiffres-clés obtenus *via* le logiciel CR Viewer (§1). Cette sous-section a pour objectif de montrer, de manière globale, les phénomènes des CR représentés dans notre corpus.

À partir de ces données, nous expliciterons les emplois des marqueurs référentiels et leurs caractéristiques du point de vue inter-langues (§2). Pour ce faire, nous examinerons les contextes d'occurrence des pronoms, des noms propres et des SN. Nous analyserons également la particularité et la tendance d'utilisation des expressions référentielles qui se présentent différemment selon la langue.

Enfin, dans la dernière sous-section, nous évoquerons l'effet du genre textuel (§3). Avant de débiter notre discussion sur les romans, nous montrerons quelques extraits de textes encyclopédiques pour tester rapidement si la même tendance d'utilisation des marqueurs référentiels est observée dans les textes non-narratifs. Ensuite, les traits typiques des textes narratifs et la particularité importante de notre corpus seront décrits afin de donner une vue synthétique sur l'influence du genre textuel et sur les CR.

1 Analyse quantitative générale

À l'aide de CR Viewer¹, nous avons récupéré les données quantitatives globales relatives aux CR dans les deux textes. En complément, quelques calculs manuels seront faits afin de décrire plus précisément à la fois les caractéristiques du genre textuel et de chaque langue.

1.1 Aperçu sur les chiffres-clés

Notre corpus comprend deux textes inter-langues dont chacun se compose des trois premiers chapitres du roman. Ces chapitres totalisent environ 10 000 mots dans les deux cas. Le tableau suivant récapitule les chiffres-clés de notre corpus, obtenus *via* le CR Viewer :

Paramètres	Corpus français : <i>Fantôme</i> ²	Corpus coréen : <i>Arbre</i> ³
Nombre de chapitres	3 (Les deux premiers chapitres et 2/3 du troisième chapitre)	3
Nombre de mots	10 422	9 743
Nombre de maillons	1 985	1 945
Densité référentielle ⁴	19 %	19,9 %
Nombre de chaînes	161	184
Taille (longueur) moyenne des chaînes	12,3 maillons	10,6 maillons
Distance inter-maillonnaire moyenne	498,9 mots	623 mots
Longueur moyenne des maillons	2,1 mots	1,6 mots

¹ Le logiciel fournit certaines mesures sur tout ou partie du corpus ou bien sur certaines chaînes particulières (voir le sous-thème 3 dans le chapitre 5).

² L'abréviation pour le roman français *Le Fantôme de l'Opéra*.

³ L'abréviation pour le roman coréen *Arbre à feuillages persistants*.

⁴ Calculée en divisant le nombre d'expressions référentielles (ici, le nombre de maillons) par le nombre de mots du texte.

Densité en CR ⁵	1,5 %	1,9 %
----------------------------	-------	-------

Tableau 10 Tableau récapitulatif des chiffres-clés des deux corpus

Le corpus français comprend 10 422 mots dont 1 985 expressions référentielles (19 %) participant des CR dont le nombre total s'élève à 161. La densité en chaînes, calculée par le nombre des chaînes sur le nombre de mots du texte entier est donc de 1,5 %. Chaque chaîne comprend en moyenne 12,3 maillons constitués approximativement de 2,1 mots et un maillon apparaît en moyenne tous les 498,9 mots, ce qu'on appelle la distance inter-maillonnaire.

Le corpus coréen, quant à lui, comprend 9 743 mots, dont 1 945 marqueurs référentiels (ce qui donne 19,9% de densité référentielle) qui participent à la constitution de 184 chaînes. La densité en chaîne, évaluée à 1,9 %, est légèrement supérieure à celle du français, du fait de la plus petite taille du texte coréen par rapport à celle du texte français. En moyenne, 10,6 maillons constituent une chaîne et ces maillons contiennent approximativement 1,6 mots. La distance inter-maillonnaire est de 623 mots.

Les chiffres importants ayant été présentés succinctement, il est à présent nécessaire d'examiner plus précisément ce qui est révélé par ces indicateurs. Une approche quantitative présente des limites dans le sens où elle n'est pas en mesure de fournir une explication approfondie sur les CR. En effet, nous ne remarquons que de légères différences inter-langues *via* ces données générales et celles-ci n'évoquent pas la corrélation entre les CR et le texte. De ce fait, dans ce qui suit, nous prendrons comme point de départ de notre analyse la répartition des catégories grammaticales afin d'appréhender plus finement les caractéristiques des CR et des expressions référentielles dans les deux textes écrits en deux langues différentes.

1.2 Corpus français

Les maillons de *Fantôme* se décomposent ainsi :

⁵ La proportion calculée par le nombre des chaînes sur le nombre de mots du texte.

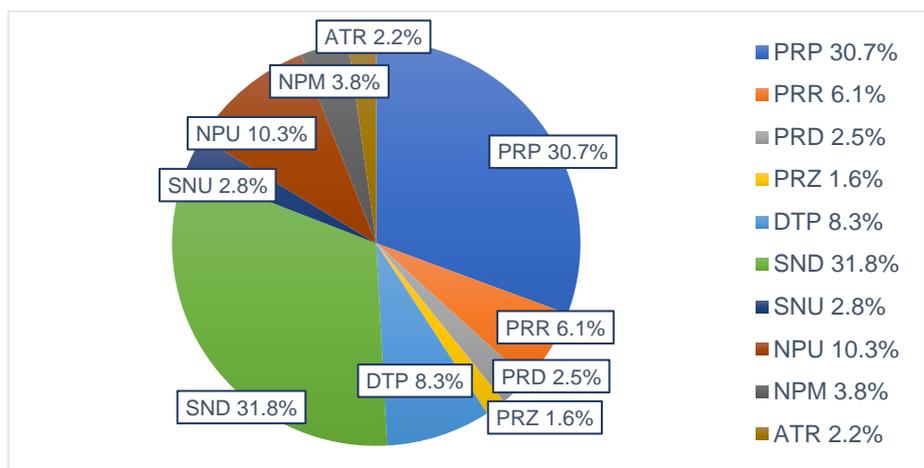


Figure 27 Répartition des catégories grammaticales des maillons⁶ (*Fantôme*)

Les pronoms (40.9 % des expressions référentielles) et les SN anaphoriques (34,6 %) occupent les deux premières places au niveau de la catégorie grammaticale avec, respectivement, 811 et 687 occurrences.

PRO	40,9 % (811 occ.)	Pronom personnel : 30,7 %
		Pronom démonstratif : 2,5 %
		Pronom relatif : 6,1 %
		Pronom zéro : 1,6 %
SN	34,6 % (687 occ.)	SN avec déterminant : 31,8 %
		SN nu : 2,8 %
NPR	14,1 % (279 occ.)	Npr nu : 10,3 %
		Npr modifié : 3,8 %
DetP	Déterminant possessif 8,3 % (165 occ.)	
AUTRES	Autres 2,2 % (43 occ.)	

Tableau 11 Les occurrences des catégories grammaticales (*Fantôme*)

Dans ce texte, à titre indépendant, le SN avec déterminant est la forme la plus fréquente de reprise nominale (31,8 %), suivi par le pronom personnel (30,7 %).

⁶ Dans les figures, nous utiliserons les abréviations proposées dans la section 3.3 du chapitre 5. (PRP : Pronom personnel / PRR : Pronom relatif / PRD : Pronom démonstratif / PRZ : Pronom zéro / DTP : Déterminant possessif / SND : SN déterminé / SNU : SN nu / NPU : Nom propre nu / NPM : Nom propre modifié / ATR : Autres).

1.3 Corpus coréen

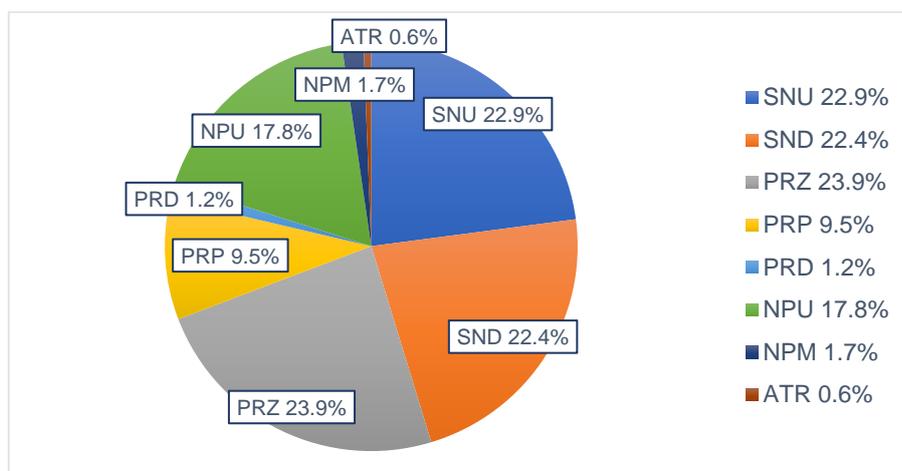


Figure 28 Répartition des catégories grammaticales des maillons (*Arbre*)

À la différence du corpus français, dans le corpus coréen, les SN anaphoriques sont prédominants avec 45,3 % des expressions référentielles, suivis par les pronoms (34,6 % des expressions référentielles). Cependant, dans le détail, on observe une grande différence entre les deux corpus : en coréen, le pronom personnel (9,5 %) est utilisé environ trois fois moins qu'en français (30,7 %), mais le pourcentage du pronom zéro est beaucoup plus élevé : 23,9 % dans le texte coréen et 1,6 % dans le texte français⁷.

SN	45,3 % (882 occ.)	SN nu : 22,9 %
		SN avec déterminant : 22,4 %
PRO	34,6 % (672 occ.)	Pronom personnel : 9,5 %
		Pronom démonstratif : 1,2 %
		Pronom zéro : 23,9 %
NPR	19,5 % (380 occ.)	Npr nu : 17,8 %
		Npr modifié : 1,7 %
AUTRES	Autres 0,6 % (11 occ.)	

Tableau 12 Les occurrences des catégories grammaticales (*Arbre*)

Les deux langues manifestent un certain nombre de disparités dans la répartition des catégories grammaticales. Dans la section suivante, nous examinerons ces différences et expliciterons les caractéristiques de chaque marqueur référentiel et leur contexte d'occurrence dans chacun des deux textes⁸.

⁷ Voir le tableau 11 *supra*.

⁸ Nous avons préféré analyser plus particulièrement la composition des CR plutôt que de tester de nombreux

2 Les différences entre les langues dans la composition des chaînes

Comme nous venons de le mentionner sur la base des tableaux 11 et 12, des disparités importantes émergent entre le français et le coréen. Le tableau suivant récapitule les données obtenues *via* les deux tableaux précités et hiérarchise les catégories grammaticales employées dans chaque langue :

	Fr : <i>Fantôme</i>			Co : <i>Arbre</i>		
Hiérarchie des composants	PRO > SN > NPR (40,9%) (34,6%) (14,1%)			SN > PRO > NPR (45,3%) (34,6%) (19,5%)		
Trois catégories prédominantes	SND 31,8%	PRP 30,7%	NPU 10,3%	PRZ 23,9%	SNU 22,9%	SND 22,4%

Tableau 13 La hiérarchie des composants des CR dans les deux langues

Dans les pages suivantes, nous nous intéresserons à l'utilisation des expressions référentielles dans chaque langue en nous basant sur des exemples extraits de notre corpus dans le but de trouver des explications à cette disparité proportionnelle.

2.1 L'interférence potentielle et le choix des expressions référentielles

Ariel (1990) souligne qu'une expression référentielle indique le degré de saillance du référent⁹. Ainsi, l'emploi du pronom, qui signale une haute accessibilité, signifie que le référent est plus saillant que celui désigné par le nom propre.

Toutefois, la présence de plusieurs référents saillants ayant le même nombre et le même genre grammatical risque de causer des ambiguïtés. C'est pourquoi, dans le but d'explicitier cette sorte de situations référentielles, Givon (1983) présente la notion d'*interférence*

paramètres. Ce choix peut se justifier par deux aspects. Premièrement, l'un des objectifs de cette thèse est de comparer l'emploi des expressions référentielles dans les deux langues. Deuxièmement, nous envisageons d'étudier des CR à long terme ce qui explique la raison pour laquelle nous avons constitué un corpus d'un genre textuel relativement long par rapport à ceux traités par nos prédécesseurs (voir la section 1.1 du chapitre 5). Notre intérêt se porte donc sur une perspective d'enchaînement textuel, c'est-à-dire les CR dans des segments linguistiques au-delà du niveau intra-phrastique ou intra-paragraphique. En conséquence, les paramètres concernant la dimension intra-phrastique, comme par exemple la fonction syntaxique de maillon dans la phrase ou intra-paragraphique tel que le nombre de maillons/chaînes par paragraphe ne seront pas étudiés.

⁹ Voir la section 2.1 du chapitre 2.

potentielle. L'auteur insiste sur le fait que « la zone de compétition "dangereuse" se situerait en deçà de trois propositions » (Schneidecker, 2021 : 103). Les exemples (1) et (2) qui suivent illustrent ce phénomène :

(1) [*Fantôme* #139]

Ce soir-là, la confusion n'avait jamais été plus complète, mais **Raoul** n'avait jamais été moins timide. **Il** écartait d'une épaule solide tout ce qui **lui** faisait obstacle, ne s'occupant point de ce qui se disait autour de **lui**, n'essayant point de comprendre les propos effarés des machinistes. **Il** était uniquement préoccupé du désir de voir celle dont la voix magique **lui** avait arraché le cœur.

Dans cet exemple, *Raoul* est mentionné et repris par les pronoms personnels « il » et « lui ». Le sujet de cette série de phrases, *Raoul*, reste saillant tout au long du texte, sans aucun risque de créer d'interférence avec un autre référent identique en genre et en nombre, ce qui justifie l'emploi fréquent de pronoms personnels.

En revanche, lorsque la concurrence se fait entre plusieurs référents partageant le même trait de genre et de nombre, l'utilisation du pronom personnel peut augmenter le coût cognitif pour interpréter la référence. Dans de tels cas, on peut choisir le *nom propre* ou le *SN* afin d'éliminer le risque d'ambiguïté, comme dans l'exemple suivant (2) :

(2) [*Fantôme* #118]

Christine Daaé avait été jusqu'alors un Siebel convenable auprès de cette Marguerite un peu trop splendidement matérielle qu'était *la Carlotta*. Et il avait fallu l'absence incompréhensible et inexplicable de *la Carlotta*, à cette soirée de gala, pour qu'au pied levé **la petite Daaé** pût donner toute **sa** mesure dans une partie du programme réservée à *la diva espagnole* ! Enfin, comment, privés de *Carlotta*, MM. Debienne et Poligny s'étaient-ils adressés à **la Daaé** ? Ils connaissaient donc **son** génie caché ? Et s'ils le connaissaient, pourquoi le cachaient-ils ? Et **elle**, pourquoi le cachait-**elle** ? Chose bizarre, on ne **lui** connaissait point de professeur actuel. **Elle** avait déclaré à plusieurs reprises que, désormais, **elle** travaillerait toute seule. Tout cela était bien inexplicable.

Dans cet exemple, les deux chaînes de *Christine Daaé* et de *la Carlotta* s'entrecroisent dans les trois premières propositions : les maillons de chaque chaîne s'alternent au fur et à mesure du déroulement du texte. À chaque fois, le narrateur utilise les expressions

référentielles de basse accessibilité pour éviter l'interférence potentielle de l'une à l'autre¹⁰. Les quatre premiers maillons de la chaîne *Christine Daaé* sont « Christine Daaé... la petite Daaé... sa... la Daaé » et ceux de la chaîne *la Carlotta* sont « la Carlotta... la Carlotta... la diva espagnole... Carlotta ». À partir de la quatrième proposition où il n'y a plus de concurrence référentielle en cotexte local, le référent *Christine Daaé* commence à être repris par les pronoms personnels et le déterminant possessif. Nous retrouvons dans Combettes (1986) également souligné par Schnedecker (1997 : 32) la description de ce phénomène comme suit :

Tant que l'émetteur utilise un pronom, c'est que pour lui, tout se déroule normalement pourrait-on dire. Le retour, dans la ligne du texte, au groupe nominal est le signe, la trace d'un problème (Combettes 1986 : 81).

En envisageant sous cet angle les données fournies dans les tableaux 11, 12 et 13 *supra* nous nous faisons face à un dilemme puisque la proportion totale des marqueurs pleins (les noms propres et les SN) en coréen est plus importante qu'en français. Ce texte coréen, ferait-il face à davantage d'interférences potentielles, ce qui ferait choisir plus d'expressions référentielles de basse accessibilité qu'en français ? Les CR dans une langue comme le coréen, qui a tendance à utiliser moins de pronoms et plus d'expressions référentielles de basse accessibilité, demandent-elles moins d'efforts cognitifs que celles du français, dans le sens où les expressions de ce type contiennent déjà la description nécessaire du référent, pour interpréter la référence ? Notre objectif dans cette section est d'apporter des éléments de réponse à ces différentes questions.

Regardons l'exemple (3)¹¹ :

(3) [*Arbre*, de #447 à #449] : la chaîne « Dong-hyuk »

¹⁰ Le risque d'ambiguïté référentielle peut être interprété différemment selon l'unité textuelle que nous traitons : lorsque nous traitons le texte entier, autrement dit à la lecture globale du texte, ce risque provoqué par *la Carlotta* est atténué. Pourtant, pour le besoin de démontrer le phénomène d'interférence référentielle, nous avons choisi cet extrait qui peut amplifier l'ambiguïté.

¹¹ Par commodité, dans les cas où nous devons montrer la structure globale du texte, nous reproduisons les exemples en coréen avec leur traduction en français (que nous avons nous-même traduits) sans translittération. Cette traduction est faite mot à mot, ce qui induit que cela répercute l'absence d'article de la langue coréenne. Dans le texte coréen, à côté des expressions référentielles en question, nous marquons la traduction littérale en français entre parenthèses afin de bien guider le lecteur.

1_[DH] 박첨지 양주의 눈이 부시도록 넘어가는 석양을 등뒤에 받으면서 잿배기를 넘어오는 동혁 (**Dong-hyuk descendant d'une butte, suivi du soleil couchant piquant les yeux de M et Mme Park**)의 윤곽은 점점 뚜렷이 나타났다. 2_[DH] 회색 저고리 바지에 검정 조끼를 입고 삼을 둘러멘 동혁이 (**Dong-hyuk portant un pantalon gris et une veste noire avec une pelle sur l'épaule**)는 역광선에 원체 건장한 체격이 더한층 걸대가 커 보인다.

3_[DH] 아들 (**ils**)이 가까이 오자, "점심두 안 들어와 먹구 여태 어디서 뭇들을 4_[DH] (**Ø**) 했니 (**as fait**)?" 하고 묻는 아버지의 목소리는 아까 꾸짖던 때와는 뉘뉘로 부드럽다.

La silhouette de 1_[DH] **Dong-hyuk descendant d'une butte, suivi du soleil couchant qui pique les yeux de M et Mme Park**, devient de plus en plus claire. 2_[DH] **Dong-hyuk portant un pantalon gris et une veste noire avec une pelle sur l'épaule** semble plus grand qu'il est à cause du contre-jour.

Dès que 3_[DH] **le fils** s'est approché, le ton du père est devenu doux contrairement à ce qu'il était plus tôt quand il était en train de se plaindre : "4_[DH] (**Ø**) **as fait** quoi sans déjeuner jusqu'à cette heure-là ? "

Dans cet extrait, la chaîne « Dong-hyuk » comprend 4 maillons : deux NPR modifiés (1^{er} et 2^{ème} maillons), un SN nu (3^{ème} maillon) et un pronom zéro (4^{ème} maillon) dans un discours rapporté¹². Toutefois, étant donné que cette partie se consacre à la description de l'apparition de Dong-hyuk et que les maillons sont contigus, il suffirait alors tout simplement d'employer le pronom personnel pour les 2^{ème} et 3^{ème} maillons. Cependant, comme nous pouvons l'observer dans ce texte, des expressions référentielles de basse accessibilité, autrement dit les marqueurs pleins, sont employées. Dans l'exemple ci-dessous, nous essayons donc de substituer le nom propre « Dong-hyuk » du 2^{ème} maillon et « fils » du 3^{ème} maillon au pronom personnel « il » afin de savoir s'il s'agit d'un choix dû à une contrainte syntaxique ou non :

(3') 1_[DH] 박첨지 양주의 눈이 부시도록 넘어가는 석양을 등뒤에 받으면서 잿배기를 넘어오는 동혁 (**Dong-hyuk descendant d'une butte, suivi du soleil couchant piquant les yeux de M et Mme Park**)의 윤곽은 점점 뚜렷이 나타났다. 2_[DH] 회색 저고리 바지에 검정 조끼를 입고 삼을 둘러멘 그 (**il**) portant un pantalon gris et une veste noire avec

¹² Nous ne faisons pas la différence entre les références présentées dans la narration et les discours rapportés des personnages.

une pelle sur l'épaule))는 역광선에 원체 건장한 체격이 더한층 걸대가 커 보인다.

3_[DH] 그 (il)가 가까이 오자, "점심두 안 들어와 먹구 여태 어디서 뭇들을 4_[DH] (Ø) 했니 (as fait)]?" 하고 묻는 아버지의 목소리는 아까 꾸짖던 때와는 땀판으로 부드럽다.

La silhouette de 1_[DH] Dong-hyuk descendant d'une butte, suivi du soleil couchant qui pique les yeux de M et Mme Park], devient de plus en plus claire. 2_[DH] *il portant¹³ un pantalon gris et une veste noire avec une pelle sur l'épaule] semble plus grand qu'il est à cause du contre-jour.

Dès qu' 3_[DH] il s'est approché, le ton du père est devenu doux contrairement à ce qu'il était plus tôt quand il était en train de se plaindre : "4_[DH] (Ø) as fait] quoi sans déjeuner jusqu'à cette heure-là ? "

Compte tenu du fait que la substitution du nom propre « Dong-hyuk » par « il » dans le 2^{ème} maillon ne pose aucun problème, nous pouvons, pour le moment, faire le constat suivant : dans le texte coréen, bien que l'on puisse employer le pronom personnel à la place du nom propre sans contrainte, ni syntaxique ni référentielle, le nom propre est choisi. Il en va de même pour le SN comme nous venons de le voir dans le cas de « fils » du 3^{ème} maillon de (3').

Nous allons à présent nous pencher sur la description des usages des marqueurs référentiels dans nos textes afin d'examiner si le phénomène présenté *via* l'exemple (3) pourrait être considéré comme une tendance générale de la langue coréenne ou bien s'il s'agit d'emplois exceptionnels dus à une condition spécifique du discours ; si tel est le cas, nous souhaitons alors mettre cet élément en évidence dans cette analyse.

2.2 Usages des pronoms dans les deux langues

L'exemple (3) vu précédemment montre à la fois que l'emploi des marqueurs pleins est moins limité et que les pronoms sont moins utilisés en coréen qu'en français. Se pose ainsi une question essentielle : le pronom a-t-il une valeur différente en coréen ?

Pour répondre à cette question, nous expliciterons les particularités concernant les usages des pronoms dans les deux langues. Commençons par le tableau suivant qui montre

¹³ La construction « il portant un pantalon... » en français serait syntaxiquement incorrecte alors qu'en coréen, cette construction ne pose pas de problème.

la répartition des pronoms dans les deux textes :

	PRP Pronom personnel	PRR Pronom relatif	PRD Pronom démonstratif	PRZ Pronom zéro
FR : <i>Fantôme</i> (811 occ. au total)	75,1 % (609 occ.)	14,9 % (121 occ.)	6,2 % (50 occ.)	3,8 % (31 occ.)
CO : <i>Arbre</i> (672 occ. au total)	27,4 % (184 occ.)	-	3,6 % (24 occ.)	69 % (464 occ.)

Tableau 14 Répartition des pronoms utilisés dans le corpus

À part le fait que le pronom relatif (PRR) n'existe pas en coréen, les deux textes se distinguent considérablement par la proportion de pronoms personnels (PRP) et de pronoms zéro (PRZ) comme mentionné plus haut au point 1.3. En effet, le pronom personnel est environ trois fois plus utilisé en français qu'en coréen alors que le pronom zéro est à peu près dix-huit fois moins employé. La proportion de pronoms démonstratifs (PRD) est approximativement deux fois plus grande en français (2,5%) qu'en coréen (1,3%). Ainsi la fréquence d'apparition des pronoms zéros en coréen est loin d'être anodine alors que l'usage des autres types de pronoms est relativement moins important.

2.2.1 Les contextes d'apparition du pronom zéro

En coréen, le pronom zéro¹⁴ prend la première place parmi toutes les catégories grammaticales et il occupe plus de la moitié de la proportion des pronoms employés (69 %). L'exemple (4) montre le contexte où l'emploi du pronom zéro joue un rôle important :

(4) [*Arbre*, #402 et #404] : la chaîne « mère de YS » dans le discours direct de Young-shin

"제 사정은 백선생밖에는 아무한테도 말한 적이 없어요. 1[Mère de YS **홀로** 되신 우리 어머니 (**ma mère devenue veuve**)]는, 2[Mère de YS **육십 노인 (dame âgée de 60 ans)**]이 딸 하나 공부를 시키느라구 입때 생선 광주리를 이고 땡기세요. **올 여름엔** 3[Mère de YS (**Ø**) **더위를 잡숫고 (a eu chaud)**] 길바닥에 가 4[Mère de YS (**Ø**) **쓰러지신 걸 (s'est évanouie)**], 동네 사람들이 5[Mère de YS (**Ø**) **업어 다가 (a mis sur dos)**] 6[Mère de YS (**Ø**)

¹⁴ Le pronom zéro en coréen ne fonctionne pas de la même manière que celui du français et n'est pas catégorisé de la même manière dans la linguistique coréenne (voir la section 3.3.4 du chapitre 4). Pour une comparaison inter-langues, nous avons uniquement inclus les actions coordonnées dans l'ordre chronologique. Le cas de la simultanéité, souvent exprimé par le gérondif en français et par le pronom zéro en coréen, n'est pas annoté.

눅혀(a couché)] 드렸어요. 그럴건만 약 한 첩 번번히...

“Je n’ai jamais parlé de ma situation sauf à Mme Baek. Quant à 1[Mère de YS **ma mère qui est devenue veuve**], c’est 2[Mère de YS **une dame âgée de 60 ans**]¹⁵ et (elle¹⁶) vend toujours des poissons pour financer les études de (sa) fille unique. Cet été, 3[Mère de YS (Ø) **a eu trop chaud**] et 4[Mère de YS (Ø) **s’est évanouie**] dans la rue, les voisins 5[Mère de YS (Ø) **’ont portée**] et 6[Mère de YS (Ø) **’ont couchée**]. Même après, ce n’était pas possible d’acheter un médicament...

7[Mère de YS (Ø) **정신을 잃으신 (a perdu conscience)**] 동안에 어느 몫쓸놈이 8[Mère de YS (Ø) **푼푼이 모아 넣으신 (a mis pièce à pièce)**] 돈주머니를 끌러 가서 그게 9[Mère de YS (Ø) **원통해(triste)**] 밤새두룩 10[Mère de YS (Ø) **우시는데 (a pleuré)**]...!"

Pendant qu’7[Mère de YS (Ø) **avait perdu conscience**], un voyou a pris le porte-monnaie dans lequel 8[Mère de YS (Ø) **a mis pièce à pièce**], ça 9[Mère de YS (Ø) **’a rendue tellement triste**] qu’10[Mère de YS (Ø) **a pleuré**] toute la nuit... !"

Dans cet exemple, on introduit une anecdote sur *la mère de Young-shin*. Ainsi, les deux premières phrases fournissent des informations sur les circonstances entourant le personnage en question, mais, à partir de la troisième proposition « Cet été... », on commence à introduire une anecdote à propos de ce personnage. Autrement dit, dès le moment où l’on perçoit le thème principal de l’énonciation et où l’on réalise le fait que le thème (qui est bien le sujet) du discours continue, on commence à employer le pronom zéro. Nous pouvons donc en déduire qu’en coréen, le pronom zéro est utilisé avec moins de restrictions comparativement au français. Regardons l’exemple (4’) où sont indiqués en gras tous les pronoms personnels qui devraient être marqués si le texte était écrit en français, mais qui sont réalisés en tant que pronoms zéros en coréen :

(4’) “Je n’ai jamais parlé de ma situation sauf à Mme Baek. Quant à ma mère qui est devenue veuve, c’est une dame âgée de 60 ans et (**elle**) vend toujours des poissons

¹⁵ Dans une telle construction, normalement on annote « c’ » (voir la section 2.3.6.2 du Ch.6). Pourtant, dans la version coréenne de cet extrait, il n’y a pas d’équivalent de « c’est ». Nous avons employé « c’est » pour une raison de facilité de traduction, il n’est donc pas annoté ici.

¹⁶ L’expression mise entre parenthèses dans la traduction française signifie qu’elle n’est pas directement mentionnée dans la version coréenne, mais répercutée sur la traduction pour faciliter la compréhension. Il en va de même pour l’exemple (4’) *infra*.

pour financer les études de **(sa)** fille unique. Cet été, **elle** a eu trop chaud et s'est évanouie dans la rue, les voisins l'ont portée et l'ont couchée. Même après, ce n'était pas possible d'acheter un médicament...

Pendant qu'**elle** avait perdu conscience, un voyou a pris le porte-monnaie dans lequel **elle** a mis pièce à pièce, ça l'a rendue tellement triste qu'**elle** a pleuré toute la nuit... !"

Nous en remarquons ainsi neuf qui incluent les sept déjà marqués en tant que maillons dans (4). Si, dans le texte coréen, on écrivait le SN soit « ma mère » soit « mère » ou bien encore qu'on mettait le pronom personnel à la place de chaque pronom zéro, ce texte semblerait incohérent, voire même agrammatical. Les seules possibilités de restitution¹⁷ sont les deux pronoms surlignés dans (4') qui sont sujet du premier verbe de leurs phrases respectives.

2.2.2 Usage restreint du pronom personnel en coréen

D'après ce que nous venons de voir, nous pouvons estimer qu'en coréen, dans de nombreux cas, le pronom zéro se manifeste même aux endroits où le pronom personnel s'emploierait en français. Cela mis à part, un autre point important concernant le pronom de la troisième personne est à noter. En effet, en coréen, ce pronom a commencé à être usité uniquement à l'écrit à partir du début du 20^{ème} siècle ; cette restriction d'usage semble perdurer jusqu'à aujourd'hui¹⁸. Toutefois, de nombreux débats portent sur le point d'introduction précis de ce pronom dans la langue coréenne. Pour notre part, nous suivons l'idée la plus répandue selon laquelle la distinction masculin / féminin du pronom personnel de la troisième personne se réalise petit à petit à partir du début du 20^{ème} siècle pour des besoins de traduction de la littérature anglaise (Kim Kyung-suk, 2010 : 105). De ce fait, cette catégorie grammaticale a donc un statut différent des pronoms des première et deuxième personnes et nous trouvons une particularité concernant son emploi dans notre roman. Examinons les deux exemples suivants afin d'illustrer notre propos :

(5) [*Arbre*, #136] : la chaîne de « Young-shin »

¹⁷ Nous avons choisi ici, le terme *restitution* afin de désigner la substitution du pronom zéro à l'expression « visible » comme « elle ».

¹⁸ Cf. An So Jin (2008).

동혁은 여러 학생들 틈에 섞여서 서대문행 전차를 탔다. 전차가 마약 떠나려는데, 1[ys 놓치면 큰일이나 날 듯이 뛰어오르는 한 여학생 (**une étudiante sautant comme si ce serait catastrophique de rater**)] 이 있다.

2[ys 그 (il)]는 3[ys 동혁에게 생후 처음으로 깊은 인상을 준 채영신 (**Young-shin Chae ayant impressionné Dong-hyuk pour la première fois**)]이었다.

Dong-hyuk est monté dans le tram direction Seodaemun avec d'autres étudiants. Le tram est sur le point de départ et il y a 1[ys **une étudiante qui saute dans le tram comme si ce serait une catastrophe si elle le ratait**].

2[ys C]'est 3[ys **Young-shin Chae qui a impressionné Dong-hyuk pour la première fois**].

(6) [Arbre, de #241 à 246] : la chaîne « Mme Baek »

영신이가 입에 침이 말라서 동혁의 외모와 행동을 그려 내니까 1[Mme Baek **백씨 (Mme Baek)**]는, "오우 그래? 온 저런. 매우 좋은 청년이로군." 하고 서양 여자처럼 연방 감탄사를 늘어놓았다.

Comme Young-shin décrivait l'apparence et le comportement de Dong-hyuk sans arrêt, 1[Mme Baek **Mme Baek**] a énuméré des phrases exclamatives comme (si elle était) une femme d'occident.

« Ah c'est vrai ? Oui. Un bon garçon, quoi. »

2[Mme Baek **그 (il)**]는 팔베개를 하고 자리 위에 비스듬히 누워 결눈질로 흘끔흘끔 영신의 3[Mme Baek **(Ø) 눈치를 살피더니**], "아—니, 영신이가 대번에 그 남자한테 훌쩍 반한 게 아냐?" 하고 거침없이 4[Mme Baek **(Ø) 한마디를 하고**] 사내처럼 꺾꺾꺾 5[Mme Baek **(Ø) 웃는다**].

2[Mme Baek **Elle**] s'est couchée sur le côté et 3[Mme Baek **(Ø) a regardé Young-shin de côté**]. « Mais, Young-shin est amoureuse de cet homme ? » 4[Mme Baek **(Ø) a dit**] sans aucune hésitation et 5[Mme Baek **(Ø) a ri**] comme un garçon.

Dans la chaîne de *Young-shin* de l'exemple (5), le pronom personnel de la troisième personne au masculin « il » est employé en tant que deuxième maillon. Il en va de même pour la chaîne de *Mme Baek* dans (6). Il est ici important de faire remarquer que le pronom personnel de la troisième personne féminine équivalant à « elle », « geu nyeo », existe bien

en coréen mais que celui-ci n'est pas communément utilisé, et qu'il est d'ailleurs introuvable dans cette œuvre. Généralement, la reprise est effectuée avec le pronom masculin « *geu (il)* » et ce quel que soit le genre du référent. Il est important de garder en tête que le coréen est une langue qui, à l'origine, n'a pas de système de genre, la présence de « *geu (il)* » à la place de « *geu nyeo (elle)* » est donc totalement acceptable, voire l'emploi de « *geu nyeo* » semblerait, dans certains cas, manquer de naturel.

2.3 Usages des noms propres dans les deux langues

Comme nous l'avons vu *supra* dans le tableau 13, la proportion des noms propres s'élève à 14,1 % en français et à 19,5 % en coréen. La composition détaillée de cette catégorie (cf. tableau 15 ci-dessous) révèle deux différences importantes entre le français et le coréen.

	NPU Nom propre nu	NPM Nom propre modifié	
FR : <i>Fantôme</i> (279 occ. au total)	73,1 % (204 occ.)	26,9 % (75 occ.)	Longueur moyenne des maillons : 3,4 mots
CO : <i>Arbre</i> (380 occ. au total)	91,3 % (347 occ.)	8,7 % (33 occ.)	Longueur moyenne des maillons : 4,8 mots

Tableau 15 La comparaison des sous-catégories de nom propres dans les deux corpus

Nous souhaitons à présent aborder la question de l'emploi du nom propre nu (NPU), dont la proportion est supérieure en coréen, et du nom propre modifié (NPM) qui est, proportionnellement parlant, environ trois fois plus souvent observé en français qu'en coréen.

2.3.1 Les contextes d'apparition du nom propre non modifié

Comme l'illustre le tableau *supra*, le nom propre nu, dont la proportion totale s'élève à 91,3 %, représente la majorité des noms propres usités dans le texte coréen, tandis qu'en français cette catégorie occupe près des trois-quarts de tous les noms propres employés. Dans cette partie, nous examinerons les conditions d'apparition de cette catégorie grammaticale représentées dans notre corpus.

Dans l'exemple suivant, deux chaînes, mises en gras, dont le référent est donné au moins une fois par le nom propre (soit par son prénom, soit par son nom de famille ou par son nom complet) s'entrecroisent :

(7) [Fantôme, de #143 à #147]

Ainsi, [Médecin *le médecin*] et 1[Raoul **l'amoureux**] se trouvèrent dans le même moment aux côtés de [Christine **Christine**], [Christine **qui**] reçut les premiers soins de l'un et ouvrit les yeux dans les bras de 2[Raoul **l'autre**]. *Le comte* était resté, avec beaucoup d'autres, sur le seuil de la porte devant laquelle on s'étouffait.

« Ne trouvez-[Médecin *vous*] point, [Médecin *docteur*], que ces messieurs devraient "dégager" un peu la loge ? demanda 3[Raoul **Raoul**] avec une incroyable audace. [Christine **On**] ne peut plus respirer ici.

– Mais 4[Raoul **vous** avez parfaitement raison », acquiesça [Médecin *le docteur*], et [Médecin *il*] mit tout le monde à la porte, à l'exception de 5[Raoul **Raoul**] et de [Femme de chambre *la femme de chambre*].

[Femme de chambre *Celle-ci*] regardait 6[Raoul **Raoul**] avec des yeux agrandis par le plus sincère ahurissement. [Femme de chambre *Elle*] ne 7[Raoul **il**] n'avait jamais vu.

[Femme de chambre *Elle*] n'osa pas toutefois 8[Raoul **le**] questionner.

La chaîne de *Christine* et celle de *Raoul* sont respectivement constituées de diverses expressions référentielles : le SN déterminé (*l'amoureux*), le déterminant sans nom (*l'autre*), le nom propre nu (*Christine* et *Raoul*), le pronom relatif (*qui*) et le pronom personnel (*on*, *vous* et *le*). Parmi ces expressions, les noms propres nus et le SN sont utilisés dans les contextes où plusieurs personnages se présentent dans la même scène et dans les cas de prise de parole afin de distinguer les locuteurs¹⁹. Dans de telles situations, la reprise par le nom propre, entre autres, peut contribuer à l'entretien de saillance du référent principal, parce que théoriquement, comme le souligne Schnedecker (1997 : 73), « moins un référent serait saillant, plus il aurait besoin d'être redénoté ». Par exemple, à la place du sixième maillon de la chaîne *Raoul*, on aurait pu utiliser « celui-ci ». Pourtant, l'emploi de *Raoul* dans ce contexte renvoie clairement au bon référent en évitant le risque d'interprétation erronée qui pourrait être provoquée par la proposition précédente.

L'exemple (8) est un paragraphe du texte coréen décrivant le moment où le personnage *Dong-hyuk* réalise qu'il pleut :

(8) [Arbre, #513] : la chaîne de « Dong-hyuk »

¹⁹ Voir Capin *et al.* (2021 : 100-102).

1_[DH] **동혁(Dong-hyuk)** 은 덧문을 밀쳤다. 습기를 축축히 머금은 밤바람이 방 안으로 휘돌아 들자, 자던 얼굴에 방울방울 부딪히는 찬 빗방울의 감촉! 2_[DH] **동혁(Dong-hyuk)** 은 정신이 번쩍 들었다.

1_[DH] **Dong-hyuk** pousse la porte extérieure. La tornade du vent du soir humide dans la salle et la goutte de pluie froide touchant la tête qui avait sommeil ! 2_[DH] **Dong-hyuk** se réveille tout d'un coup.

Tenant compte de l'absence de la concurrence référentielle et de celle du dialogue dans le contexte, il semble que le deuxième maillon peut être remplacé par des expressions référentielles de haute accessibilité : le pronom personnel ou le pronom zéro peuvent être employés au lieu du nom propre comme dans (8'a) et (8'b) :

(8') a. **동혁 (Dong-hyuk)**은 덧문을 밀쳤다. 습기를 축축히 머금은 밤바람이 방 안으로 휘돌아 들자, 자던 얼굴에 방울방울 부딪히는 찬 빗방울의 감촉! **그 (II)**는 정신이 번쩍 들었다.

Dong-hyuk pousse la porte extérieure. La tornade du vent du soir humide dans la salle et la goutte de pluie froide touchant la tête qui avait sommeil ! **II** se réveille tout d'un coup.

b. **동혁 (Dong-hyuk)**은 덧문을 밀쳤다. 습기를 축축히 머금은 밤바람이 방 안으로 휘돌아 들자, 자던 얼굴에 방울방울 부딪히는 찬 빗방울의 감촉! **(∅)** 정신이 번쩍 들었다.

Dong-hyuk pousse la porte extérieure. La tornade du vent du soir humide dans la salle et la goutte de pluie froide touchant la tête qui avait sommeil ! **(∅)** se réveille tout d'un coup.

Les trois cas sont tous acceptables, c'est-à-dire que la contrainte syntaxique qui empêche l'usage du pronom personnel ou du pronom zéro n'existe pas dans ce paragraphe. Cependant, employer ici le nom propre *Dong-hyuk*, manifeste une rupture contextuelle et produit un effet de sens qui met en avant ce moment de la réalisation du personnage, en particulier, parce que la pluie joue un rôle important dans le récit. En revanche, les deux autres options ne peuvent pas donner lieu à un tel effet dramatique.

Analysons à présent l'exemple suivant qui, de prime abord, semble similaire à l'exemple (8) mais qui pourtant présente un caractère différent :

(9) [Arbre, de #530 à #531] : la chaîne de « Dong-hyuk »

1_[DH] 동혁(Dong-hyuk)은 창 밖으로 팔을 내밀고 천금을 주고도 그 한 방울을 살 수 없는 생명수를 손바닥에 받아 본다. 자리옷을 활활 벗어 버리고 뛰어나가서 그 비에 온몸을 골고루 적시다가 땅 위에 디굴디굴 구르고 싶은 충동을 2_[DH] (∅) 느꼈다.

3_[DH] 동혁(Dong-hyuk)은 아우가 감기나 들까 보아 다시 문을 닫았다.

1_[DH] Dong-hyuk passe (son) bras par la fenêtre et reçoit dans (sa) paume l'eau de vie qu'on ne pourrait acheter même pour tout l'or du monde. 2_[DH] (∅) ressent l'envie soudaine de se déshabiller, sortir, se tremper par cette pluie et se rouler par terre.

3_[DH] Dong-hyuk, ayant peur que (son) petit frère prenne froid, referme la porte.

Dans cet extrait, le troisième maillon, le nom propre *Dong-hyuk* peut être remplacé par le pronom personnel *il* mais non par le pronom zéro. Regardons les deux versions modifiées de (9) :

(9') a. 동혁 (Dong-hyuk)은 창 밖으로 팔을 내밀고 천금을 주고도 그 한 방울을 살 수 없는 생명수를 손바닥에 받아 본다. 자리옷을 활활 벗어 버리고 뛰어나가서 그 비에 온몸을 골고루 적시다가 땅 위에 디굴디굴 구르고 싶은 충동을 (∅) 느꼈다.

그 (II)는 아우가 감기나 들까 보아 다시 문을 닫았다.

Dong-hyuk passe (son) bras par la fenêtre et reçoit dans (sa) paume l'eau de vie qu'on ne pourrait acheter même pour tout l'or du monde. (∅) ressent l'envie soudaine de se déshabiller, sortir, se tremper par cette pluie et se rouler par terre.

II, ayant peur que (son) petit frère prenne froid, referme la porte.

b. 동혁 (Dong-hyuk)은 창 밖으로 팔을 내밀고 천금을 주고도 그 한 방울을 살 수 없는 생명수를 손바닥에 받아 본다. 자리옷을 활활 벗어 버리고 뛰어나가서 그 비에 온몸을 골고루 적시다가 땅 위에 디굴디굴 구르고 싶은 충동을 (∅) 느꼈다.

*(∅) 아우가 감기나 들까 보아 다시 문을 닫았다.

Dong-hyuk passe (son) bras par la fenêtre et reçoit dans (sa) paume l'eau de vie qu'on ne pourrait acheter même pour tout l'or du monde. (∅) ressent l'envie soudaine de se déshabiller, sortir, se tremper par cette pluie et se rouler par terre.

*(∅) ayant peur que (son) petit frère prenne froid, referme la porte.

Par ce test, nous pouvons constater que la contrainte syntaxique est plus forte dans l'exemple (9). En effet, celle-ci est liée au contenu sémantique des phrases alignées c'est-à-dire l'action du personnage décrite par la dernière phrase « refermer la porte » va à l'encontre de son désir, ce qui vient d'être détaillé par les deux premières constructions. Ainsi, dans les deux exemples de (9'), une conjonction de coordination telle que « mais » devant la dernière phrase rendrait le pronom personnel « il » plus naturel ; de même, le pronom zéro serait également acceptable : *mais, il / (Ø) ayant peur que son petit frère prenne froid, referme la porte*²⁰.

Regardons un autre exemple :

(10) [Arbre, de #542 à #546] : la chaîne de « Dong-hyuk » et celle de « Kun-bae »

1_[DH] 동혁(Dong-hyuk)은 단독으로 영신을 맞아들이고 싶지 않았던 것이다.

1_[KB] 건배(Kun-bae)는, "흥, 인제야 2_[DH] 자네(Tu)가 몽달귀신을 면허나 보이. 앞으로 다섯 해 안에는 결혼을 안 한다구 3_[DH] (Ø) 장담을 허더니 허는 수 있나. 지남철 기운에 끌려오는걸."하고 연방 4_[DH] 동혁(Dong-hyuk)을 놀려 댄다.

5_[DH] 동혁(Dong-hyuk)은 변색을 하며, "여보게, 그게 무슨 가당치 않은 소린가? 아예 그런 말은 입 밖에두 2_[KB] (Ø) 내지 말게. 동지와 애인을 구별 못 허는 6_[DH] 낸(je) 줄 3_[KB] (Ø) 아나?" 하고 4_[KB] 건배(Kun-bae)의 험구를 틀어막았다.

1_[DH] Dong-hyuk] ne voulait pas accueillir Young-shin tout seul.

1_[KB] Kun-bae], en disant « Pff, là, 2_[DH] tu] ne seras plus un fantôme de jeune homme²¹. 3_[DH] (Ø) étais tellement sûr de ne pas te marier dans les prochains cinq ans mais tant pis ! C'est magnétique ! », se moque de 4_[DH] Dong-hyuk].

5_[DH] Dong-hyuk], en prenant un air sérieux, « Hé, qu'est-ce que (tu) dis ? 2_[KB] (Ø) ne dis plus un truc pareil. Crois 3_[KB] (Ø) que 6_[DH] je] ne suis pas capable de distinguer

²⁰ Schnedecker (1997 : 127-128) souligne que la redénomination immédiate peut être acceptable par le fait de «lever la contradiction, en connectant les propositions à l'aide d'un connecteur contre-argumentatif... ou d'une expression indiquant qu'il n'y a pas simultanéité entre les événements». Cependant, dans cet exemple écrit en coréen, nous observons le cas contraire : la redénomination immédiate *via* le nom propre ne demande pas l'emploi d'un connecteur alors que c'est nécessaire pour que les pronoms soient acceptables.

²¹ Expression désignant un homme qui est mort sans se marier.

l'amour de la fraternité ? » empêche 4_[KB Kun-bae] de parler.

Dans l'exemple (10), la chaîne de *Dong-hyuk* se constitue ainsi : [NPU... PRP(tu)... PRZ... NPU... NPU... PRP(je)]. Le nom propre employé en tant que 4^{ème} maillon, *Dong-hyuk*, qui est ici l'objet de la phrase, peut être remplacé soit par le pronom personnel « geu (il) » soit par le pronom zéro alors que le 5^{ème} maillon, le sujet, est plus rigide. Ce dernier peut donc être remplacé par « geu », sous certaines conditions, mais non par le pronom zéro. Pour que le 5^{ème} maillon soit remplacé par le pronom personnel, il faudrait un adverbe tel que « aussitôt » en tête de phrase susceptible de montrer la connexité immédiate entre deux séquences d'action. Dans ce dernier cas, faute d'un tel élément adverbial, la redénomination *via* le nom propre est exigée afin de distinguer le locuteur au sein du dialogue.

Ainsi, comme nous venons de le voir, la reprise par le nom propre est communément observée dans les deux langues dans le cas où la désambiguïsation du référent est exigée en raison de la présence d'un référent potentiellement concurrentiel ou des tours de paroles des personnages. Il s'agit d'un moyen d'entretenir la saillance du référent principal dans de telles situations référentielles. En dehors de ces points communs, nous avons remarqué *via* quelques exemples du texte coréen que la redénomination par le nom propre de cette langue se présente dans des cas où un effet spécifique tel que la mise en relief d'un contexte se réalise au cours de la narration et où la connexité immédiate entre les phrases alignées n'est explicitée ni par l'élément adverbial ni par les conjonctions de sorte qu'une « rupture » contextuelle apparaît. Ce dernier usage, notamment, mérite d'être souligné comme il montre une différence nette avec la reprise *via* le nom propre en français. Contrairement au français, le nom propre repris immédiatement dans le texte coréen peut indiquer la connexité des propositions contiguës sans aide des autres éléments textuels.

2.3.2 Les différents types de nom propre modifié

La proportion de noms propres modifiés du texte français est plus importante que dans le texte coréen. Elle occupe 26,9 % de tous les noms propres dans *Fantôme* contre seulement 8,7 % dans *Arbre*. À l'exception de cette différence propositionnelle, nous remarquons que la longueur moyenne des mots constituant un nom propre modifié est différente dans chaque langue : 3,4 mots pour le français et 4,8 pour le coréen. Ce décalage ne semble pas important, mais il n'est pas trivial non plus vu les différents pourcentages de types de modifications que le nom propre accompagne. Dans le texte français, c'est souvent la forme « dét+NPR » ou bien « dét+adj+NPR » qui provoque un effet spécifique dans l'interprétation :

(11) [*Fantôme*, #151]

Il était radieux. Il conclut : « C'est **un Chagny** ! » et il se dirigea vers la loge de **la Sorelli**...

Parmi les noms propres modifiés du texte « *Fantôme* », la forme simple « dét+NPR » comme « un Chagny » et « la Sorelli » dans (11) s'élève à 53 %. Quand le nom propre est modifié par l'article indéfini comme « un Chagny », ce que nous avons défini comme le type dénominatif des noms propres²², il souligne un aspect spécifique lié aux gens portant le nom « Chagny ». En revanche, le nom propre modifié par l'article défini comme « la Sorelli » cause un effet de familiarité vis-à-vis d'une figure célèbre. En règle générale, il s'agit d'une manière de désigner des personnalités célèbres, et, dans ce roman, certains personnages, notamment les chanteuses et danseuses de l'Opéra considérées comme « connues » par d'autres personnages du récit, sont souvent nommés sous la forme de « la + NPR » ou bien celle de « la petite + NPR ». L'exemple (12) est une partie reprise de l'exemple (2) *supra* :

(12) [*Fantôme*, #118]

Christine Daaé avait été jusqu'alors **un Siebel convenable** auprès de cette Marguerite un peu trop splendidement matérielle qu'était la Carlotta. Et il avait fallu l'absence incompréhensible et inexplicable de la Carlotta, à cette soirée de gala, pour qu'au pied levé **la petite Daaé** pût donner toute sa mesure dans une partie du programme réservée à la diva espagnole !

Dans cet extrait, « un Siebel convenable » réfère au rôle « Siebel ». Il ne renvoie pas au rôle lui-même, mais à la version jouée par *Christine Daaé* accompagnée du jugement du locuteur sur cette interprétation. Pourtant, « la petite Daaé », comme « la Sorelli » de (11) ci-dessus, désigne une personnalité connue. Ainsi, dans le texte *Fantôme*, les deux formes mentionnées « dét + NPR » et « dét + adj + NPR » sont majoritairement observées au sein de tous les emplois des noms propres modifiés.

En revanche, dans le texte coréen, seuls 12 % (4 occurrences parmi 33) des noms propres modifiés apparaissent sous la forme « dét + NPR »²³ : c'est le cas des expressions telles que « jeon Chosun (toute la Corée) », « i Miss Chae (cette Mademoiselle Chae) » et « i

²² Voir la section 2.3.2.1 du chapitre 4.

²³ Il s'agit des cas « *Gwanhyungsa* + NPR » (NPR employé avec les expressions démonstrative ou quantificative).

Mister Park (ce Monsieur Park)²⁴ ».

La plupart des noms propres modifiés en coréen ont souvent une structure plus compliquée. Dans le texte coréen, 75,7 % des noms propres modifiés sont expansés par des relatives. L'exemple (13) est extrait de l'exemple (3) introduit plus haut :

(13) [*Arbre*, #447]

박첨지 양주의 눈이 부시도록 넘어가는 석양을 등뒤에 받으면서 잣배기를 넘어오는 동혁 (Dong-hyuk descendant d'une butte, suivi du soleil couchant piquant les yeux de M. et Mme Park)의 윤곽은 점점 뚜렷이 나타났다. 회색 저고리 바지에 검정 조끼를 입고 삼을 둘러멘 동혁이 (Dong-hyuk portant un pantalon gris et une veste noire avec une pelle sur l'épaule)는 역광선에 원체 건장한 체격이 더한층 걸대가 커 보인다.

La silhouette de **Dong-hyuk descendant d'une butte, suivi du soleil couchant qui pique les yeux de M. et Mme Park**, devient de plus en plus claire. **Dong-hyuk portant un pantalon gris et une veste noire avec une pelle sur l'épaule** semble plus grand qu'il ne l'est à cause du contre-jour.

Les constructions surlignées équivalent à la construction relative en français et sont employées majoritairement afin de modifier le nom propre dans le texte coréen. Ces expressions ayant au moins un prédicat montrent la manière dont le coréen transmet l'information sur le référent portant le nom propre en question. Par exemple, le premier maillon de (13) décrit l'action de *Dong-hyuk* et l'arrière-plan de cette action *via* la structure relative : *Dong-hyuk* est en train de descendre d'une butte, derrière lui, on voit le soleil couchant et le soleil pique les yeux de M. et Mme Park. Ces trois différentes propositions sont liées et incluses dans une relative qui modifie le sujet *Dong-hyuk*. Bien évidemment, nous trouvons cette sorte d'expression dans le texte français comme le montre l'exemple (14) :

(14) [*Fantôme*, #189]

J'ai dit que cette fête magnifique avait été donnée, à l'occasion de leur départ de l'Opéra, par **MM. Debiegne et Poligny qui avaient voulu mourir comme nous**

²⁴ Nous considérons le nom de titre d'une personne comme une partie d'un nom propre. Il en va de même pour le français : « M Poligny » est annoté comme un seul NPR.

disons aujourd'hui : en beauté.

Pourtant, les expressions de ce type y sont beaucoup moins utilisées (9,5 %, soit 7 occurrences sur 74). En revanche, nous rencontrons parfois des cas où la structure relative est employée de façon appositive, à savoir, insérée entre virgules comme dans l'exemple suivant :

(15) [*Fantôme*, #4]

La Sorelli, **qui désirait être seule un instant pour « repasser » le compliment qu'elle devait prononcer tout à l'heure au foyer devant MM. Debienne et Poligny**, avait vu avec méchante humeur toute cette foule étourdie se ruer derrière elle.

Contrairement au français, l'apposition par une incise n'est pas grammaticale en coréen, de ce fait, nous employons à la place la structure « relative + NPR » et ce quelles que soient la longueur ou la quantité informationnelle de la relative²⁵. En d'autres termes, souvent, le coréen transmet l'information sur le référent par le biais de la structure interne du syntagme nominal. Ces exemples nous permettent donc ainsi de constater que la façon dont on modifie le nom propre varie selon la langue.

2.4 Usages des SN anaphoriques dans les deux langues

Reprenons maintenant les données présentées dans le tableau 13 *supra* (voir la section 2). Dans le texte français, 34,6 % des expressions référentielles sont des SN anaphoriques, tandis qu'ils représentent 45,3 % dans le texte coréen.

Le SN déterminé, lui, (31,8 %) est environ onze fois plus employé que le SN nu (2,8 %) ce qui s'explique par le fait que le français est une langue avec un système d'articles. En revanche, en coréen, qui ne possède pas un tel système, le SN nu a un pourcentage légèrement supérieur (22,9 %) à celui du SN déterminé (22,4%). Cette distribution indique non seulement que le SN anaphorique occupe un peu moins de la moitié de toutes les expressions référentielles mais aussi que cette moitié est elle-même divisée en deux parties : la première correspondant au nom déterminé par l'élément qualificatif avec lequel il est employé ; la deuxième, correspondant à sa détermination qui est dépendante du contexte comme expliqué dans la section 2.3 du chapitre 3.

²⁵ Voir la section 2.2.6.2 du chapitre 6.

2.4.1 Les contextes d'apparition du SN nu

En français, les SN nus se manifestent dans quelques contextes typiques. Nous évoquerons trois cas selon qu'un SN nu apparaît dans des expressions figées, qu'il joue un rôle de classifiant et qu'il occupe une fonction appellative. Les deux premiers cas ayant déjà été évoqués dans le chapitre précédent (voir respectivement la section 2.2.1 et 2.2.2.2 du chapitre 6), chaque usage sera brièvement rappelé avec les exemples du texte français.

De prime abord, le référent non actualisé qui fait partie d'une expression figée, peut être désigné par le SN nu, tel que le « nez » dans l'exemple ci-dessous :

(16) [*Fantôme*, #18]

Il s'était heurté, – on ne saurait dire « nez à **nez** », car le fantôme n'**en** avait pas, – avec le mystérieux personnage dans le petit escalier qui, près de la rampe, descend directement au « dessous ».

Dans l'exemple (16), le « nez » est employé dans une expression adverbiale figée, il n'a donc normalement pas le statut de référent actualisé. Cependant, le terme étant repris par le pronom « en », il devient annotable et, dans ce cas-là, la reprise provoque une sorte de défigement et de jeu de mot²⁶.

Regardons à présent l'exemple (17) qui illustre le deuxième cas de figure, le SN nu ayant un rôle de classifiant :

(17) [*Fantôme* #3 et #118]

Ce soir-là, qui était celui où MM. Debienne et Poligny, les directeurs démissionnaires de l'Opéra, donnaient [Dernier Gala leur dernière soirée de 1[GALA **gala**]], à l'occasion de leur départ...

Et il avait fallu l'absence incompréhensible et inexplicable de la Carlotta, à [Dernier Gala cette soirée de 2[GALA **gala**]], pour qu'au pied levé la petite Daaé pût donner toute sa mesure dans une partie du programme réservée à la diva espagnole !

Dans les SN « leur dernière soirée de gala » et « cette soirée de gala » de (17), le terme « gala » n'est pas actualisé. Dans le SN de forme « dét. + N1 de N2 », le N2 joue un rôle

²⁶ Ce type d'emploi du SN nu est, théoriquement, tout à fait possible en coréen. Pourtant, nous n'avons pas trouvé d'exemples précis dans notre corpus.

classifiant analogue à celui de l'adjectif.

Ensuite, les termes désignant le métier ou le statut d'une personne s'emploient sous la forme de SN nu lorsqu'ils sont vocatifs. C'est le cas du premier maillon de l'exemple (18), « docteur » dans le discours de *Raoul* :

(18) [*Fantôme* #144 et #145]

« Ne trouvez-vous point, **docteur**, que ces messieurs devraient “dégager” un peu la loge ? demanda Raoul avec une incroyable audace. On ne peut plus respirer ici.

– Mais vous avez parfaitement raison », acquiesça **le docteur**, et **il** mit tout le monde à la porte, à l'exception de Raoul et de la femme de chambre.

Les termes basés sur les relations interpersonnelles sont souvent inclus dans ce type d'usage comme le montre l'exemple suivant :

(19) [*Fantôme* de #156 à #158]

« **Mademoiselle**, puisqu'il vous plaît de ne point me reconnaître, je voudrais vous dire quelque chose en particulier, quelque chose de très important.

– Quand j'irai mieux, monsieur, voulez-vous ?... – et sa voix tremblait. – Vous êtes très gentil...

– Mais il faut vous en aller... ajouta le docteur avec son plus aimable sourire. Laissez-moi soigner **mademoiselle**.

Les appellations « mademoiselle », « monsieur » utilisés dans les dialogues entre personnages ne s'emploient pas avec le déterminant²⁷. L'emploi de « Maman » est aussi fréquemment observé dans notre corpus.

Dans le texte coréen, nous avons remarqué que les SN nus sont employés non seulement dans des contextes typiques que nous venons de mentionner mais aussi dans des situations plus étendues. L'exemple (20) *infra* présente deux sortes d'usage de SN nu :

(20) [*Arbre*, #13 et #16]

²⁷ Parmi ces termes mentionnés, « mademoiselle » a un statut particulier par rapport aux autres. On dira moins facilement « une mademoiselle » qu'« un monsieur ».

"긴 말씀은 허지 않겠으나, 1_[Thé] 차(thé)나(-na/particule auxiliaire)] 마셔 가면서 간담적으로 피차에 의견도 교환하고, 그 동안에 분투한 체험담도 들려주셔서 앞으로 이 운동을 계속하는 데 크게 참고가 되게 해주시기를 바라는 바입니다."

« (Je) ne vais pas détailler, (je) voudrais que (vous) échangiez librement vos avis en prenant 1_[Thé] (au moins une boisson comme) thé], (vous) (nous) aidez à continuer ce mouvement en racontant vos histoires vécues. »

...L여학교 가사과의 학생들은, 굉장한 연회나 차리는 듯이 일제히 에이프런을 두르고 돌아다니며 자기네의 손으로 만든 과자와 2_[Thé] 차(thé)]를 주욱 돌린다.

« Les étudiantes en éducation maternelle du lycée L faisant un tour avec leurs tabliers, comme si c'était un grand banquet, distribuent (aux autres) les biscuits faits maison et 2_[Thé] du thé. »

Dans ce contexte, le premier maillon « thé » est un référent non actualisé employé en jouant un rôle de classifiant, à savoir, l'adjectif. La particule auxiliaire « -na » s'emploie après le nom en ajoutant le sens de la modestie, c'est-à-dire que cette particule dévalorise le nom antéposé et donne l'impression que le référent de ce nom n'a pas d'importance. Dans ce texte, « 차나 (thé-na) » signifie donc que *nous n'avons pas pu préparer grand-chose, mais nous vous offrons « (au moins une boisson comme du) thé »*. Le sens générique du « thé » est ici utilisé pour désigner de manière modeste la boisson préparée. En revanche, le deuxième « thé » est un référent concret. Dans les deux cas, ils sont exprimés au moyen de SN nus.

Maintenant, regardons l'exemple (21) qui concerne la fonction appellative d'un SN nu :

(21) [Arbre, de #467 et #470] : la chaîne de Dong-hyuk

"아 그래, 1_[DH] 성님(**grand frère**)]은 공부두 혼자 하구 밥꺼정 혼자 먹는 거유?"

...아우의 성미를 건드렸다가는 마구 뚫린 창구멍으로 무슨 소리가 나올지 몰라서

2_[DH] 형(**grand frère**)]은 점잖이 타이른다.

« Et alors, 1_[DH] **grand frère**] (tu) étudies tout seul et manges

tout seul ? »

...2_[DH] **grand frère**] ayant peur de ce qui va sortir de la bouche de son petit frère, le dit gentiment.

Dans cet exemple, le premier « grand frère » dans le discours du petit frère renvoyant à

Dong-hyuk est un mot basé sur la relation interpersonnelle. L'emploi vocatif des termes de ce type, sous la forme de SN nu, est également observé dans le texte coréen. Cependant, comme le montre la deuxième occurrence de « grand frère », le SN nu est utilisé même dans une circonstance narrative, autrement dit, sans fonction vocative, sans relation interpersonnelle des personnes concernées.

Ainsi, les caractéristiques du coréen, qui ne possède pas de système d'articles et qui s'appuie sur d'autres éléments du contexte pour sa détermination, ont des répercussions sur l'emploi des SN. À la différence du français, les SN nus en coréen s'emploient non seulement dans des circonstances typiques, à savoir les trois cas évoqués au début de cette partie, mais aussi dans des situations plus étendues comme le deuxième « thé » de l'exemple (18) ou le deuxième « grand frère » de l'exemple (21).

2.4.2 Les différents types de SN anaphorique

Selon l'échelle d'accessibilité d'Ariel (1990), les SN définis et les démonstratifs sont supposés renvoyer à des référents de moyenne accessibilité. L'auteur distingue les SN longs et courts au sein de cette catégorie²⁸ : le SN long, dont le nom est expansé par une relative, porte plus d'informations sur le référent en question et ce dernier est censé avoir une moindre accessibilité que les SN courts de types « le + N » et « ce + N ». En outre, nous voulons souligner que ces deux catégories de SN courts peuvent être distinguées encore une fois par leurs différents niveaux d'accessibilité. Ariel (1990) met ainsi en exergue que « les marqueurs de haute accessibilité exigent moins d'efforts tandis que les marqueurs de basse accessibilité exigent plus d'efforts pour récupérer l'antécédent à partir de mémoire²⁹ » (*ibid* : 53). Sous cet angle, nous distinguons donc le SN défini du SN démonstratif : le premier nécessite davantage d'efforts pour récupérer l'antécédent que le deuxième qui renvoie immédiatement à un référent plus proche. Le tableau suivant revisite alors la catégorisation d'Ariel sur les SN anaphoriques en nuancant les différents niveaux d'accessibilité :

²⁸ Ariel (1990 : 53) : « En fait, une fois que le locuteur ajoute plus d'information, par exemple, sous la forme d'une construction relative, les marqueurs de l'accessibilité médiane se transforment en marqueurs bien qualifiés de basse accessibilité (In fact, once the speaker adds more information, for instance in the form of a relative clause, Intermediate Accessibility Markers turn into full-fledged Low accessibility Markers) » (traduit par nos soins).

²⁹ « High accessibility markers imply minimal effort, while low accessibility markers imply greater efforts in recovering the antecedent from memory » (traduit par nos soins).

	Basse accessibilité	Description définie	longue ³⁰	Dét + N + relative
			brève	Dét + A + N / Dét + N de N
				Dét + N
	Haute accessibilité	Expression démonstrative	longue	Ce + N + relative
			brève	Ce + A + N
				Ce + N

Tableau 16 La catégorisation des SN anaphoriques selon le degré d'accessibilité

À partir de ces critères, nous avons calculé la proportion de chaque type d'expression pour analyser la composition globale des SN anaphoriques de chaque langue. Examinons à présent la composition des SN anaphoriques dans les deux textes dans le tableau suivant :

	Fantôme	Arbre
Description définie	90,1% de tous les SN (563 occ.)	93,6 % (827 occ.)
	Longue : 6,8 %	Longue : 18,2 %
	Brève : 82,4 %	Brève : 75,4 %
Expression démonstrative	9,9% (68 occ.)	6,3 % (56 occ.)
	Longue : 0,5 %	Longue : 0,6 %
	Brève : 10,3 %	Brève : 5,8 %

Tableau 17 La composition des SN anaphoriques des deux textes

Ce tableau indique quelques disparités entre les deux textes. Comme nous venons de l'observer dans l'emploi du nom propre modifié souvent accompagné d'une relative (voir la section 2.3.2 *supra*), nous remarquons que dans le texte coréen la description définie longue, à savoir les noms expansés par une construction relative, se présente 2,7 fois plus fréquemment que dans le texte français. Regardons les trois premiers maillons de la chaîne « militants » de l'exemple (22) :

(22) [Arbre, #3 et #4] : la chaîne « militants »

가을 학기가 되자, 1[Militants ○○일보사에서 주최하는 학생계몽운동에 참가하였던 대원들 (militants ayant participé au mouvement d'éclaircissement organisé par JournalOO)] 이 돌아왔다. 오늘 저녁은 2[Militants 각처에서 모여든 대원들 (militants venant de partout)] 을 위로하는 다과회가 그 신문사 누상에서 열린 것이다.

³⁰ Ce terme ne convient pas exactement à la langue coréenne dans laquelle il n'y a pas d'article. Pourtant nous l'utilisons en tant que tel par commodité. La notion désigne le SN ayant une forme de « déterminant (ou même sans déterminant pour certains cas en coréen) +N + relative ».

A la rentrée scolaire, 1[Militants **les militants qui avaient participé au mouvement d'éclaircissement organisé par le Journal OO**] sont rentrés. Ce soir, une réunion pour encourager 2[Militants **les militants venant (qui sont venus) de partout**] a lieu dans le bâtiment de ce Journal.

오륙백 명이나 수용할 수 있는 대강당에는 3[Militants **전 조선의 방방곡곡으로 흩어져서 한여름 동안 땀을 흘려 가며 활동한 남녀 대원들 (militants homme et femmes ayant travaillé en transpirant tout l'été partout dans le pays Chosun)**]로 빈틈없이 들어찼다.

L'amphi susceptible d'accueillir de 500 à 600 personnes, est complètement rempli par 3[Militants **des militants et militantes qui ont travaillé dur pendant tout l'été dans toutes les régions de Corée**].

En revanche, la description définie brève du français, évaluée à 82,4 %, est supérieure à celle du texte coréen (75,4 %) avec 7 % d'écart. Concernant l'emploi des expressions démonstratives longues, les deux textes montrent une similitude au niveau de la proportion (0,5 % pour le français et 0,6 % pour le coréen) tandis que le français tend à utiliser 1,8 fois plus souvent la forme simple que le coréen.

Ces chiffres démontrent encore une fois, la propension du coréen à utiliser davantage, et ce proportionnellement au français, les expressions référentielles de basse accessibilité.

2.5 Les points de fraction et les marqueurs référentiels

Concernant le contexte d'occurrence de certaines expressions référentielles, Schnedecker (1997) présente la notion de *changement situationnel* qui coïncide souvent avec la redénomination. Les points de fraction, à savoir la « rupture » contextuelle comme la modification des paramètres spatiaux, temporels, énonciatifs ou autres dans le discours sont signalés par l'occurrence des marqueurs pleins.

Selon les précisions de Baumer (2015)³¹ qui retient l'idée de Schnedecker (1997) pour analyser son corpus, les ruptures aspectuo-temporelles sont représentées par des adverbes

³¹ Dans son étude, il examine également la relation entre le découpage paragraphique et la redénomination. Il s'agit d'une manière efficace d'analyser la structuration textuelle qui mérite d'être appliquée. Cependant, nous ne pouvons pas l'appliquer à notre étude, parce que le roman *Arbre* étant publié en feuilleton dans le journal ne suit pas le découpage paragraphique au sens moderne.

temporels ou des changements de temps grammatical et d'aspect, les ruptures spatiales qui sont des repérages spatiaux dans le récit sont marquées par des groupes à valeur adverbiale, les changements de plan d'énonciation se présentent souvent par des commentaires du narrateur sur son récit et enfin les ruptures argumentatives sont souvent effectuées par un connecteur contre-argumentatif ou par un changement soudain de thème. Selon son analyse sur les types de rupture et les occurrences de marqueurs pleins dans les nouvelles françaises, la majorité des noms propres et des SN anaphoriques (qui ne coïncident pas avec le découpage de paragraphe) se présentent lors de la rupture aspectuo-temporelle.

Bien que cette approche semble intéressante, il était cependant difficile de l'appliquer à notre analyse et d'obtenir les chiffres de toutes les CR ; nos textes sélectionnés étant des romans relativement longs où de nombreuses CR de tous types de référents s'entrecroisent³². Cela aurait constitué un travail trop complexe d'examiner l'occurrence des marqueurs pleins en rapport avec le changement situationnel.

Toutefois, au cours de ce point consacré à la composition des CR dans les deux langues, nous avons pu remarquer que la rupture aspectuo-temporelle est plus difficile à détecter dans le texte coréen qu'en français du fait que le temps verbal du coréen ne se divise qu'en trois catégories, passé-présent-futur ; ainsi dans *Arbre*, le passé est le seul temps verbal utilisé dans la narration sans changement. En ce qui concerne les ruptures argumentatives nous avons déjà pu remarquer en analysant l'exemple (9) et la note 20 *supra*, qu'en coréen, lors d'une rupture argumentative, le connecteur contre-argumentatif n'est pas un facteur nécessaire pour que la redénomination *via* le nom propre se présente. D'ailleurs, au contraire, dans un tel contexte de rupture, c'est le connecteur qui renforce la connexité immédiate entre les phrases et permet d'employer les pronoms. Ainsi, la façon dont les marqueurs référentiels interagissent avec la rupture contextuelle peut varier selon la langue.

Tenant compte de ces particularités, nous escomptons pouvoir appliquer cette approche de manière exhaustive et catégoriser la corrélation entre les types de rupture et la redénomination dans nos futures recherches.

Au sein de cette deuxième section, nous avons montré la distribution et les contextes d'apparition des marqueurs référentiels dans les deux textes annotés. En règle générale, les

³² L'analyse de Baumer (2015) se base sur les nouvelles qui sont, en général, formellement plus simples. Elles sont relativement brèves et se déroulent autour d'un seul événement avec moins de personnages (nous aurons l'occasion d'en reparler dans la section 3.2 *infra*).

pronoms indiquent une continuité tandis que les marqueurs pleins manifestent plutôt une rupture. De ce fait, nous avons voulu comparer le français et le coréen, vu que les deux textes montrent des disparités concernant la proportion et la hiérarchie des catégories grammaticales employées. En examinant quelques exemples du corpus, nous avons remarqué que les expressions référentielles de basse accessibilité sont utilisées dans des situations référentielles potentiellement ambiguës ou dans des cas où les tours de parole des personnages se présentent, afin d'indiquer clairement le référent en question.

Cependant, dans certains cas, la langue coréenne semble moins sensible à l'utilisation des expressions référentielles de basse accessibilité que le français. En effet, dans le texte coréen, il arrive qu'un nom propre ou un SN s'emploie même dans le contexte où la présence du pronom semble tout à fait raisonnable. D'après nos observations, ce phénomène ne se présente pas de façon aléatoire, mais en fonction de deux conditions principales : tout d'abord lorsqu'un élément comme l'adverbe ou la conjonction qui est susceptible de montrer la connexité immédiate entre les propositions contiguës est absent et ensuite lorsqu'il est exigé de transmettre l'information sur le référent en question. Dans de nombreux cas, cette transmission se fait, sous la forme d'un nom propre modifié par l'intermédiaire d'une relative ou d'une description définie longue. Excepté ces deux critères d'apparition des marqueurs référentiels de basse accessibilité dans les contextes sans interférence potentielle, le « médiocre » statut de pronom personnel de la troisième personne en coréen dû à son développement relativement tardif peut également causer, en partie, une moindre utilisation de cette catégorie et davantage d'usages des marqueurs pleins.

Au terme de cette section, nous avons démontré, au moins pour partie, que l'emploi des marqueurs référentiels dans les CR peut se manifester différemment selon la langue. Dans la section suivante, nous examinerons si la façon dont les CR se présentent est tributaire de la structure textuelle et si tel est le cas, nous souhaitons explorer la corrélation entre le genre textuel et les CR.

3 Les chaînes de référence et l'influence du genre textuel

La différence inter-langue occupe une place importante dans l'analyse du fait que chaque langue a sa propre façon de structurer un texte et d'entretenir sa cohérence. Néanmoins, il

est aussi important d'examiner les traits généraux du genre textuel ainsi qu'individuels de chaque texte pour comprendre les CR. Certaines données doivent être décrites à travers les caractéristiques des textes, car la structure textuelle est aussi l'un des facteurs essentiels à considérer lors de l'analyse.

Dans ce qui suit, nous présenterons, de manière générale, les spécificités du genre narratif à considérer et la façon dont elles influencent la constitution des CR dans les deux romans de notre corpus. Cependant, avant de débiter cette présentation, il nous semble utile de nous intéresser à la tendance du coréen que nous venons de décrire dans la section précédente et qui s'observe également dans un autre genre textuel, notamment dans le texte encyclopédique.

3.1 L'influence du genre textuel : le cas du texte encyclopédique

À ce stade de notre analyse, une question se pose : toutes les différences entre les langues coréenne et française que nous venons de décrire dans la partie précédente seraient-elles valables dans d'autres genres textuels ?

Pour répondre à cette question, nous examinerons, de façon empirique, quelques extraits de textes qui ne sont ni littéraires ni narratifs mais des textes encyclopédiques (*Wikipedia*) écrits dans les deux langues.

3.1.1 Aperçu des caractéristiques du texte encyclopédique

D'après Combettes & Tomassone (1988 : 6), chaque genre possède une « intention » différente ; ainsi, celle du texte informatif est de montrer aux lecteurs les connaissances de la réalité sans essayer de les convaincre. Quant au texte encyclopédique à vocation informative, il fournit aux lecteurs des connaissances variées sur un référent.

Outre les deux règles caractéristiques proposées par Combettes & Tomassone (1988) qui sont la règle de *progression* et la règle de *répétition*³³, Longo & Todirascu (2014 : 81)

³³ D'après Combettes & Tomassone (*ibid.* : 9, cité dans Longo & Todirascu, 2014), ces deux règles du texte informatifs se définissent par :

- la règle de *progression* : « le texte doit présenter un apport d'information constamment renouvelé » ;
- la règle de *répétition* : « le lecteur doit être capable d'identifier la récurrence de certains éléments, le

récapitulent quelques caractéristiques du texte informatif :

- Hiérarchisation des informations reflétée par des divisions en sous-thèmes ;
- Utilisation fréquente d'appositions afin d'ajouter des informations supplémentaires ;
- Emploi de certains marqueurs de cohésion comme les anaphores fidèles ou infidèles et les répétitions lexicales ;
- Emploi des tournures impersonnelles, de nominalisations fréquentes, et de mots abstraits.

La classification des informations est évidente dans les textes informatifs. On traite les informations en plusieurs sous-thèmes démontrant la hiérarchisation des informations soit par le découpage paragraphique soit par le sommaire, notamment dans les textes encyclopédiques, comme cela est présenté dans la figure suivante :



Sommaire [masquer]

- 1 Histoire de la race
 - 1.1 Le XIX^e siècle
- 2 Caractère
- 3 Standard (résumé)
- 4 Soins et santé
- 5 Photos
- 6 Croisement
- 7 Notes et références
- 8 Voir aussi
- 9 Liens externes

Figure 29 Sommaire du texte sur « Caniche », Wikipédia

Schnedecker (2014 : 32) de son côté distingue encore le texte encyclopédique du texte scientifique académique en énumérant les caractéristiques formelles du premier par l'absence d'ancrage énonciatif, l'absence de « jargon » et le mode d'écriture exploitant le présent de vérité générale. Bien évidemment, même dans les textes encyclopédiques, le mode d'écriture se manifeste différemment selon « la provenance des textes ». Dans son étude, l'auteur signale quelques traits notables des textes de *Wikipedia* par rapport à d'autres textes encyclopédiques, à savoir :

même référent apparaissant d'ordinaire sous des dénominations différentes ».

[...] les textes de *Wikipedia* sont plus découpés, avec une syntaxe disons assez simple (phrases simples réduites à SVO, peu de subordination), peu de marques de connexité - et, le cas échéant, marqueurs et cadratifs temporels-, l'essentiel des liens inter-phrastiques étant dévolus (comme pour les recettes de cuisine et pour les textes de *Wikipedia*) à la cohésion référentielle. (Schneidecker, 2014 : 32).

Enfin, il est également important de prendre en compte le fait que les textes de Wikipedia sont rédigés de manière collective comme le font remarquer Oberlé *et al.* (2018) qui soulignent que « les paragraphes successifs n'ont pas forcément été rédigés par le même auteur » (2018 : 114).

3.1.2 Les chaînes de référence dans les textes encyclopédiques

Notre objectif de cette sous-section est d'examiner quelques exemples de textes encyclopédiques afin de voir rapidement si les marqueurs référentiels se présentent de la même manière que dans les romans. Pour ce faire, nous avons sélectionné cinq textes de thème différent (l'animal, le pays, le personnage, le rôle d'un poste et le sport) pour le français et le coréen. Ce choix nous aidera à voir plus clairement la corrélation entre l'usage des expressions référentielles et le type de référent dans les deux langues. Nous commencerons cette rapide analyse en testant l'utilisation du pronom.

3.1.2.1 L'emploi du pronom

Nous avons observé que dans le roman coréen, on utilise davantage de pronoms zéros et moins de pronoms personnels qu'en français. Le même phénomène se présente dans le texte informatif. Les deux exemples suivants sont des extraits de textes de *Wikipedia*. Il s'agit, dans les deux cas, de la partie introductive du texte encyclopédique sur une des races de chien. Commençons par le texte français, « Caniche » :

(23) [Extrait du texte français « Caniche », description générale]

Le caniche est une race de chien. **C'**était autrefois un chien adapté à la chasse au canard dans l'eau, ce qui explique le toilettage « en lion » qui **lui** est souvent appliqué, et qui était fort à la mode à la Belle Époque, **il** descend du barbet.

Régulièrement utilisé dans différents domaines, comme animal de cirque, chien de

chasse, chien-guide d'aveugle, chien d'agility ou chien de recherche, **il** est très polyvalent et **est** classé parmi les trois races canines les plus intelligentes.

Dans cet exemple, la chaîne *caniche* est constituée ainsi : « SN (Le caniche) ... PRD (Ce)... PRP (lui)... PRP(il)... PRP(il)... PRZ (Ø)... ». Introduit par le SN déterminé « le caniche », le référent est réinstancié par différents types de pronoms.

L'exemple suivant est issu du texte en coréen et qui correspond au même type de partie introductive de *Wikipedia* sur la race de chien, « Jindo » :

(24) [Extrait du texte coréen « **진돗개 (Jindo)** », description générale]

진돗개는 개 품종 중 하나로, 한국의 대표적인 개 품종으로 꼽힌다. (Ø) 용맹하고 충성심이 강해 주인을 잘 따른다. 그러나 (Ø) 한 주인에게 매이는 성격이 너무 강한지라 사역견으로는 못 쓴다. (Ø) 어미 (mère)³⁴는 모성애가 매우 강하다. 털은 주로 노란 것과 흰 것이 있으며, 그 외 재구, 네눈박이, 호구, 흑구 등 10여 종류의 다양한 것이 있다.

(Ø) 귀가 쫓긋 서 있고 꼬리는 위로 말려 있다. 그리고 꼬리 및 몸 전체 단모이다. 대한민국에서는 1962년 **진도 특산인 진돗개**를 천연기념물 제53호로 지정해 보호하고 있다.

Jindo est une race de chien qui est un des chiens représentatifs en Corée. Brave et loyal, (II) obéit à son maître. Mais, étant trop attaché à (son) maître, (II) n'est pas adapté à être un chien de travail³⁵. (sa) maman a une forte affection maternelle. Quant à la couleur du poil, les deux types les plus connus sont le fauve rouge et le blanc. À part eux, il y a environ 10 types comprenant le cendré, le tigré, le noir-marron, le noir, etc.

(Ils) ont les oreilles droites et la queue attachée haut et enroulée. (Leur) corps y compris la queue est couverte de poil court. En Corée, **Jindo qui est une spécialité**

³⁴ Cela peut être considéré comme une anaphore associative. Le coréen ne possédant pas de déterminant possessif, la relation métonymique est souvent exprimée par les SN « N -ui + N », « N + N » ou bien « (Ø) + N ». Ici, le terme « mère » est l'anaphorique de « Jindo » (en français, « sa mère » serait plus naturel). Toutefois, l'anaphore associative ne fait pas l'objet de l'annotation (voir la section 2.2 du chapitre 6) et nous considérons donc ici ce type de cas comme le pronom zéro.

³⁵ Comme son nom l'indique, il s'agit d'un terme générique pour désigner le chien éduqué et utilisé pour une tâche spécifique.

de la région Jindo est nommé, en 1962, comme monument naturel numéro 53 et (le) protégé.

Dans (24), la chaîne *Jindo* comprend le SN nu, le SN modifié par une relative, et le pronom zéro : « SN(Jindo)... PRZ(Ø)... PRZ(Ø)... PRZ(Ø)... SN (Jindo qui est une spécialité de la région Jindo)... ». Par rapport à la chaîne « caniche » de (23), ici, l'emploi du pronom est moins diversifié.

Voici à présent deux autres exemples qui évoquent le statut économique d'un pays dans le monde. Il s'agit également de la partie introductrice des textes sur la « France » et sur la « Corée du sud » :

(25) [Extrait du texte français « France », description générale]

La France est, en juillet 2018, la deuxième économie européenne derrière l'Allemagne, la septième économie mondiale par PIB nominal derrière l'Inde et, en avril 2016, la neuvième économie mondiale par PIB à parité de pouvoir d'achat. Malgré un taux de chômage relativement fort, **elle** affiche un niveau de vie « très élevé » (20e au classement IDH en 2014). **Elle** figure parmi les leaders mondiaux dans les secteurs de l'agroalimentaire, de l'aéronautique, de l'automobile, des produits de luxe, du tourisme et du nucléaire.

(26) [Extrait du texte coréen « 대한민국 (Corée du sud) », description générale]

대한민국(Corée du sud)은 한국 전쟁 이래 일명 '한강의 기적'이라고 불리는 높은 경제 발전을 이룩하며, 1990년대에 이르러 세계적인 경제 강국으로 발전하였다. (Ø) 2015년 구매력 기준 1인당 국민 총소득(GDP)은 36,601 달러로 세계은행에서 고소득 국가로 분류되었고, 2016년 유엔의 인간 개발 지수(HDI) 조사에서 (Ø) 세계 18위로 '매우 높음'으로 분류되었다. 또한 국제통화기금 (IMF)에서는 **대한민국(Corée du sud)**을 선진경제국으로 분류하고 있다.

La Corée du sud, en prenant son essor économique qui s'appelle « le miracle de la rivière Han » depuis la Guerre Coréenne, est devenue, dans les années 1990, une des puissances économiques mondiales. En 2015, (Ø) est classée par la Banque Mondiale parmi les pays riches par 36 601 dollars du PIB à parité de pouvoir d'achat, et (Ø) s'impose au 18^e au classement IDH de l'ONU en 2016 avec un niveau de vie « très élevé ». Le FMI classe **la Corée du sud** comme pays riche.

Dans (25), « la France » est reprise par le pronom personnel « elle », alors que ce n'est pas le cas dans le texte coréen sur le pays. Dans ce dernier, on observe l'emploi du pronom zéro et la redénomination par le nom propre. Ce phénomène correspond à ce que nous avons remarqué comme particularité lors de l'analyse du roman coréen. Pourtant, un autre point important est à noter dans ce cas précis : le coréen, contrairement au français, n'a pas de système de genre au niveau du lexique. En d'autres termes, le vocabulaire renvoyant à des entités non-humaines n'est jamais repris ni par « il » ni par « elle ». Les chaînes *jindo* et *Corée du sud* de (24) et de (26) renvoient aux référents non-humains et montrent que le pronom personnel de la troisième personne ne s'emploie pas.

3.1.2.2 L'emploi des marqueurs référentiels de basse accessibilité

En analysant l'utilisation de chaque marqueur référentiel dans les romans, nous avons pu remarquer qu'en coréen, il arrive que les marqueurs référentiels de basse accessibilité s'emploient à la place du pronom personnel sous certaines conditions. Afin de voir si le même phénomène se manifeste dans un autre genre textuel, nous examinerons d'abord le cas du référent humain dans un extrait de texte encyclopédique :

(27) [Extrait du texte coréen « 미카 (Mika) », Jeunesse]

미카 (**Mika**)는 레바논의 베이루트에서 태어났다. (Ø) 아버지는 미국인이며 (Ø) 어머니는 레바논인. 5남매 중 셋째로, (Ø) 위로는 두 명의 누나가 있으며 밑으로 여동생과 남동생이 있다. 미카 (**Mika**)가 한 살이었을 때 그리스도교와 이슬람교의 내전이 격화되어 미카 (**Mika**)의 가족은 레바논을 떠나 파리로 이주해야 했다. 7살 때 너바나의 〈Heart-Shaped Box〉를 듣고 (Ø) 첫 번째 곡을 쓰는데, 미카 (**Mika**)는 그것을 '분노'라는 제목의 꿈쩍한 피아노 반주곡이었다고 회상한다.

Mika est né à Beyrouth, Liban. (Ø) père est américain et (Ø) mère est libanaise. Étant le troisième enfant de la famille, (Ø) a deux sœurs aînées, une sœur et un frère plus jeunes. Quand **Mika** a un an, suite à l'aggravation de la guerre civile entre les chrétiens et musulmans, la famille de **Mika** part du Liban pour Paris. **Mika** se souvient de la première chanson qu' (Ø) a composée à l'âge de 7 ans, après avoir écouté « Heart-Shaped Box » de Nirvana : c'était un accompagnement au piano intitulé « la colère ».

그 (ii)가 9살 때, 아버지가 인질이 되어 쿠웨이트의 미국 대사관에 감금되는 일이

발생했고 이 사건을 계기로 영국의 비호를 받게 된 **미카 (Mika)** 일가는 런던으로 이주했다. 런던에서 **미카 (Mika)**는 프랑스인 학교에 입학했지만 동급생들의 심한 따돌림에 난독증까지 겹쳐 **(Ø)** 정상적인 생활이 어려웠다. **그 (il)**가 11살 때 교육기관에서는 **그 (il)**에게 맞는 새 학교를 찾을 때까지 **그 (il)**를 수업에서 격리시켰는데 이 기간 동안 **미카 (Mika)**는 어머니에게서 6~8개월간 자택교육을 받았고, **(Ø)** 러시아인 소프라노 오페라 가수에게 노래 지도를 받으며 타고난 재능을 연마했다.

Quand **il** a 9 ans, (son) père est enfermé à l'ambassade des États-Unis au Koweït en tant qu'otage, et en raison de cet incident, la famille de **Mika** est protégée par l'Angleterre et déménage à Londres. À Londres, **Mika** entre à l'école française mais **(Ø)** souffre de harcèlement scolaire et de problèmes de dyslexie. Quand **il** a 11 ans, l'école décide de l'isoler de ses camarades et de **lui** trouver une autre école, pendant cette période de 6-8 mois, **Mika** est éduqué par (sa) mère à la maison et **(Ø)** développe (son) talent par l'enseignement d'un chanteur opéra russe.

Dans cet exemple décrivant la vie de *Mika*, un chanteur célèbre, le pronom personnel de la troisième personne est employé même si la fréquence est plus basse que son emploi dans le texte français. Dans l'exemple (28), relatif au texte français sur la même personnalité, nous remarquons que le pronom personnel est fréquemment utilisé, alors que nous ne trouvons aucune occurrence d'un pronom zéro :

(28) [Extrait du texte français « Mika (chanteur) », Jeunesse et débuts]

Né à Beyrouth le 18 août 1983 d'une mère libanaise (née aux États-Unis et dont le père est syrien, originaire de Damas) et d'un père américain (né à Jérusalem), **Michael Holbrook Penniman Junior** a un an lorsque **sa** famille quitte le Liban pour Paris, où **il** va à l'école maternelle. **Il** vit dans la capitale française pendant 8 ans et ensuite **part** vivre à Londres, au Royaume-Uni. **Mika** a deux sœurs aînées, Yasmine (DaWack) et Paloma. **Il** a également une sœur plus jeune, Zuleika et un frère plus jeune, Fortuné. Yasmine l'aide avec le visuel de **ses** œuvres (c'est elle qui a dessiné les pochettes de **ses** albums *Life in Cartoon Motion* et *The Boy Who Knew Too Much*). Lorsqu'**il** a huit ans, **son** père est retenu à l'ambassade des États-Unis au Koweït pour des raisons de sécurité, pendant sept mois, au tout début de la Guerre du Golfe.

Revenons à la question de la fréquence d'utilisation du pronom personnel dans l'exemple (27). Dans ce texte, le pronom personnel de la troisième personne s'emploie moins souvent que les autres catégories grammaticales. Le nom propre est observé sept fois, le pronom zéro

cinq fois et le pronom personnel de la troisième personne quatre fois.

Dans (27') qui suit, nous allons substituer les trois derniers « Mika » du premier paragraphe de (27) au pronom personnel « il » afin de tester s'il s'agit d'un choix dû à une contrainte syntaxique :

(27')미카 (Mika)는 레바논의 베이루트에서 태어났다. (Ø) 아버지는 미국인이며 (Ø) 어머니는 레바논인. 5남매 중 셋째로, (Ø) 위로는 두 명의 누나가 있으며 밑으로 여동생과 남동생이 있다. 그 (il)가 한 살이었을 때 그리스도교와 이슬람교의 내전이 격화되어 그 (il)의 가족은 레바논을 떠나 파리로 이주해야 했다. 7살 때 너바나의 〈Heart-Shaped Box〉를 듣고 (Ø) 첫 번째 곡을 쓰는데, 그 (il)는 그것을 '분노'라는 제목의 꿈쩍한 피아노 반주곡이었다고 회상한다.

Mika est né à Beyrouth, Liban. (Ø) père est américain et (Ø) mère est libanaise. Étant le troisième enfant de la famille, (Ø) a deux sœurs aînées, une sœur et un frère plus jeunes. Quand il a un an, suite à l'aggravation de la guerre civile entre chrétiens et musulmans, sa famille part du Liban pour Paris. Il se souvient de la première chanson qu' (Ø) a composée à l'âge de 7 ans, après avoir écouté « Heart-Shaped Box » de Nirvana : c'était un accompagnement au piano intitulé « la colère ».

La substitution du nom propre au pronom personnel ne pose aucun problème dans ce contexte où l'on est en train d'énumérer par ordre chronologique des informations sur le référent en question. La chaîne *Mika* ne rencontre pas d'autres référents susceptibles de provoquer une interférence potentielle et les phrases contiguës ne présentent pas une corrélation évidente entre elles qui pourrait être justifiée par l'utilisation d'une conjonction ou d'un adverbe. L'exemple (27') montre qu'il n'y a pas de contrainte syntaxique empêchant l'emploi du pronom personnel de la troisième personne. Pourtant, dans (27), nous observons *la redénomination immédiate* (Schnecker, 1997)³⁶. Ainsi, le nom propre « Mika » se manifeste dans les phrases contiguës et présente à chaque fois un nouvel incident qui débute tels que le déménagement à l'étranger, la composition de la première chanson ou l'entrée à l'école. Cette reprise immédiate ne pose au contraire aucun problème, et produit un effet de sens qui indique une rupture contextuelle par le commencement d'un nouvel épisode au sein de la progression du texte.

La redénomination par le SN est également observée dans le texte coréen comme nous

³⁶ Voir la note 20 *supra*.

pouvons l'observer dans l'exemple (29) où il est question du pouvoir exécutif en Corée du sud :

(29) [Extrait du texte coréen « 대한민국 (Corée du sud) », pouvoir exécutif]

대한민국의 행정부에 해당하는 대한민국 정부는 **대통령 (président)**을 수반으로 하여 입법부에서 법률로써 정한 사안들을 실행한다. **대통령(président)**은 국무총리와 국무위원으로 구성된 국무회의의 조력을 받아 업무를 처리한다.

Le gouvernement de la Corée qui est le pouvoir exécutif, avec **le président** en tête, exerce les décisions prises par le pouvoir législatif en tant que lois. **Le président** effectue son travail à l'aide du Premier ministre et des membres du Conseil des ministres.

대통령(président)은 5년 단임으로 국민의 보통, 평등, 직접, 비밀 선거에 의하여 선출된다. **대통령(président)**은 국회가 통과한 법률을 거부할 권한이 있다. 그러나 한번 거부한 법률을 국회가 다시 통과시킨다면(단, 국회는 재적의원 과반수의 출석과 3분의 2의 동의를 얻어야 한다.) 그 법률은 그대로 통과된다.

Le président est élu pour 5 ans d'un seul mandat au suffrage universel direct au scrutin secret. **Le président** a le pouvoir de refuser la loi que le Parlement a passé. Mais si le Parlement repasse cette loi refusée (il faut néanmoins la présence de plus de la moitié des députés entiers et le consentement de leurs deux tiers), cela doit être accepté.

Afin de réaliser une comparaison juste, nous avons extrait la partie équivalente, le pouvoir exécutif en France, dans le texte de *Wikipedia* « la France » :

(30) [Extrait du texte français « France », pouvoir exécutif]

Le pouvoir exécutif appartient en premier lieu **au président de la République**, élu pour cinq ans au suffrage universel direct au scrutin uninominal majoritaire à deux tours. **Le président de la République est le chef de l'État et le chef des armées, il promulgue les lois et peut dissoudre l'Assemblée nationale. Il nomme le Premier ministre, et, sur proposition de celui-ci, les membres du gouvernement. Le gouvernement peut être renversé par une motion de censure adoptée par l'Assemblée nationale. Lorsque la majorité parlementaire et le président n'appartiennent pas au même parti politique, on parle alors de cohabitation.**

Nous trouvons la chaîne « le président » dans les deux textes qui décrivent le rôle de président, nous pouvons donc en déduire que le référent a un caractère générique. En coréen, le terme « 대통령 (président) » est répété alors que dans le texte français, le même référent est désigné par diverses expressions référentielles telles que « le président de la République », « le chef de l'État », « le chef des armées »³⁷, « il » et « le président ». Dans les exemples (29'a-b), nous substituons le deuxième *président* de chaque paragraphe au pronom personnel « il » :

(29'a) 대한민국의 행정부에 해당하는 대한민국 정부는 **대통령 (président)**을 수반으로 하여 입법부에서 법률로써 정한 사안들을 실행한다. ?**그(ii)**는 국무총리와 국무위원으로 구성된 국무회의의 조력을 받아 업무를 처리한다.

Le gouvernement de la Corée qui est le pouvoir exécutif, avec **le président** en tête, exerce les décisions prises par le pouvoir législatif en tant que lois. **Il** effectue son travail à l'aide du Premier ministre et des membres du Conseil des ministres.

(29'b) **대통령(président)**은 5년 단임으로 국민의 보통, 평등, 직접, 비밀 선거에 의하여 선출된다. **그(ii)**는 국회가 통과한 법률을 거부할 권한이 있다. 그러나 한 번 거부한 법률을 국회가 다시 통과시킨다면(단, 국회는 재적의원 과반수의 출석과 3분의 2의 동의를 얻어야 한다.) 그 법률은 그대로 통과된다.

Le président est élu pour 5 ans d'un seul mandat au suffrage universel direct au scrutin secret. **Il** a le pouvoir de refuser la loi que le Parlement a passé. Mais si le Parlement repasse cette loi refusée (il faut néanmoins la présence de plus de la moitié des députés entiers et le consentement de leurs deux tiers), cela doit être accepté.

L'emploi du pronom personnel « il » dans (29'a) est possible, mais il est moins naturel que la redénomination par le SN puisqu'il s'agit d'un référent générique. On ne désigne pas un président spécifique, mais on parle plutôt du poste de président en général³⁸. En coréen, le pronom personnel se manifeste afin de désigner une personne spécifique, à savoir un référent humain. Par conséquent, si on utilise « il » à la place de « président », les premier et deuxième maillons ne sont plus exactement coréférentiels. Afin d'éviter ce type de confusion,

³⁷ Nous avons décidé d'annoter les expressions attributives lorsqu'elles entretiennent une relation coréférentielle avec leur antécédent (voir 2.2.5 dans le chapitre 6).

³⁸ Voir la notion de *near-identity* dans la section 2.1 du chapitre 6.

dans (29), le « président » est repris à l'identique.

En revanche, le pronom personnel « il » dans (29'b) est moins problématique. En effet, les deux premières propositions parlent respectivement de l'élection et du pouvoir du président et donc ne montrent pas de connexité immédiate entre elles ce qui justifie généralement, en coréen, la reprise à l'identique. Ce qui rend l'emploi de « il » possible dans ce contexte est le prédicat du premier maillon « président », *être élu*. Il ajoute un effet de sens spécifique au référent, c'est-à-dire que le référent en question est compris comme « ce président dont on parle qui vient d'être élu... » et joue un rôle qui établit un lien thématique entre les phrases comme le font certains adverbess ou les conjonctions.

Voici des cas semblables dans les exemples (31) et (32) :

(31) Extrait du texte coréen « 피겨스케이팅 (patinage artistique) », description générale]

피겨스케이팅 (patinage artistique)은 개인, 2인조, 또는 집단이 스케이트를 신고 얼음 위를 활주하며 갖가지 동작으로 율동의 아름다움을 겨루는 경기로 '빙상의 예술'이라고도 하며, 동계 올림픽의 꽃이라고 해도 과언이 아니다. 피겨스케이팅 (patinage artistique)에서는 점프, 스핀, 리프트, 스텝, 턴 등 다양한 기술이 사용된다.

Le patinage artistique est un sport de patin à lame, appelé aussi « l'art sur glace » où on évalue la beauté des mouvements sur glace faits par un individu, un couple ou par un groupe de gens. Il n'est pas exagéré de dire qu'(il) est la fleur des Jeux olympiques d'hiver. Plusieurs figures sont employées telles que les sauts, les pirouettes, les levées, les pas, les spirales, etc. dans **le patinage artistique**.

피겨스케이팅 (patinage artistique)은 1908년에 첫 번째 동계 경기로 올림픽에 포함되었다. 올림픽 종목으로는 남녀 싱글 스케이팅, 페어 스케이팅, 아이스 댄싱의 네 가지가 있다.

Le patinage artistique a commencé à être inclus dans les Jeux olympiques d'hiver en 1908. Les compétitions comprennent quatre disciplines que sont les patinages individuels pour dames et pour hommes, le patinage en couple et la danse sur glace.

(32) [Extrait du texte français « patinage artistique », description générale]

Le patinage artistique est un sport et une discipline artistique exécutés sur la glace, synthétique ou artificielle, qui développent les qualités physiques et la sensibilité artistique. Il se pratique avec des patins à lame.

Ce sport est particulièrement populaire en Amérique du Nord, en Europe et en Asie, notamment au Japon. La Russie compte de nombreuses patineuses dans ses rangs. D'après Médiamétrie, **c'est le seul sport à être davantage regardé par les femmes que par les hommes.**

Comme nous pouvons l'observer à la lecture des deux textes informatifs sur le patinage artistique, les expressions référentielles se présentent différemment. L'exemple (31) exploite la redénomination par le SN, autrement dit, le SN « patinage artistique » est répété. Pourtant, dans (32), le référent est repris par divers types de maillons tels que « le patinage artistique », « il », « ce sport », « ce », et « le seul sport à être davantage regardé par les femmes que par les hommes ».

Par ces exemples tirés de *Wikipedia*, textes non littéraires et non narratifs, nous pouvons provisoirement conclure que contrairement au coréen, les textes encyclopédiques écrits en français cherchent la diversité lexicale des maillons. En effet, dans les textes en coréen, nous remarquons que, dans de nombreux cas, la CR du référent principal est constituée du pronom zéro et des SN ou noms propres repris à l'identique ; et, même si le référent est humain, le pronom personnel est relativement peu présent.

Bien que cette partie ait un caractère expérimental et demande donc à être approfondie, elle nous permet de voir empiriquement que la diversité lexicale des CR du coréen est moins évidente dans le texte encyclopédique que dans le roman ; de même, nous remarquons que le type de référent (humain/inhumain) joue un rôle plus essentiel dans l'utilisation des pronoms que ce à quoi nous nous attendions. De ce fait, cela indique que, même si la particularité d'une langue est un élément décisif dans l'analyse des CR, il ne faut pas négliger l'importance des caractéristiques du genre textuel. Forte de ce constat, nous nous concentrerons dans ce qui suit sur les traits du genre narratif et la structuration textuelle de notre corpus pour analyser plus précisément les CR.

3.2 État de l'art sur les chaînes de référence en contexte narratif

Comme mentionné dans la section 3.1 du chapitre 2 et la 1.1 du chapitre 3, de nombreuses recherches ont déjà été menées sur les CR et un bon nombre d'entre elles travaillent sur les textes narratifs. Pourtant, il est difficile de dégager la tendance typique d'un genre textuel, car les conditions de constitution de corpus diffèrent selon les recherches. Le tableau suivant récapitule les différents genres narratifs qui ont fait l'objet de chaque étude :

Nom de recherche	Sous-genre du corpus	Les caractéristiques des textes
Baumer (2015)	Nouvelles	40 nouvelles comprenant 7 531 mots (il y a l'équivalent en anglais).
Baumer (2017)	Nouvelles contemporaines courtes	11 nouvelles courtes totalisant 8 046 mots.
Obry, Glikman, Guillot-Barbance et Pincemin. (2017)	Récits brefs de type humoristique (fabliaux du 13 ^e siècle et nouvelles des 15 ^e et 16 ^e siècles).	2 fabliaux de 720 et 1 210 mots et 3 nouvelles constituées de 5 parties comportant chacune en moyenne 884 mots.
Schnedecker (2017)	Contes de fées	<i>Incipit</i> de 7 contes de Perrault constitués de 1 305 mots.
Baumer, Dias et Schnedecker (2020)	Roman réaliste du 19 ^e siècle	<i>Incipit</i> de 15 textes français de 7 965 mots (à peu près la même quantité en anglais et en allemand, voir le tableau 1 de l'article).
Capin, Glikman, Schnedecker et Todirascu (2021)	Texte à dominante narrative du 12 ^e au 20 ^e siècle	7 textes narratifs (sous-corpus de DEMOCRAT) comportant 86 034 mots.

Tableau 18 Tableau récapitulatif des corpus utilisés dans les recherches antérieures

Bien que ces auteurs travaillent communément sur les textes narratifs, les sous-genres, les époques de la rédaction, la taille des corpus et les objectifs d'étude sont différents. Ces différences ne sont pas triviales. Parmi ces recherches, celles de Baumer (2015 ; 2017) coïncident particulièrement avec ce qui nous intéresse parce que l'auteur travaille sur des textes français contemporains. Il caractérise les nouvelles en les comparant aux romans. En règle générale, les nouvelles sont d'un format relativement bref par rapport aux romans et centrées sur un seul événement central. Les personnages secondaires sont moins développés. En ce qui concerne « le rythme du récit » et « la gestion de la chronologie des événements », l'écriture est condensée et la fin est, dans la plupart des cas, « inattendue et surprenante » (Baumer 2015 : 37). En revanche, les romans sont relativement longs et beaucoup de personnages secondaires se présentent au sein du récit. Plusieurs événements peuvent se succéder avant d'arriver au point culminant et les perturbations sont résolues à la fin (voir la section 1.3.2 du chapitre 5).

Ainsi, chaque corpus ne se caractérise pas de la même manière, malgré son regroupement sous l'étiquette de « narratif ». Les différences entre les sous-genres, la taille des textes et les époques de la rédaction ont certaines répercussions sur les résultats. Il est donc important que nous nous servions des recherches mentionnées dans le tableau *supra*

comme guide pour notre analyse sans négliger pour autant les caractéristiques propres à nos textes sélectionnés.

Sous cet angle, cette dernière partie de notre thèse envisage l'analyse des CR et de leur corrélation avec les particularités textuelles de notre corpus *via* les paramètres mentionnés dans la section 1.1 de ce chapitre. Avant d'y procéder, il nous paraît toutefois pertinent de récapituler le contenu de chaque chapitre des deux romans de notre corpus :

Texte	Genre	Taille de texte	Contenus des textes	Annotation
FR : <i>Fantôme</i>	Roman du 20 ^e siècle	10 422 mots sur 107 317 mots du texte entier	Introduction du contexte du récit : CH1 : Présentation du cadre spatio-temporel du récit et réactions des personnages secondaires autour d'un incident fonctionnant comme point de départ de l'événement principal. Mention d'un des personnages principaux. CH2 : Entrée en scène d'autres personnages principaux au sein du même contexte spatio-temporel. Présentation des relations entre les personnages par des anecdotes. CH3 : Avènement d'un incident perturbant au sein du même contexte spatio-temporel.	Toutes les chaînes (elles sont continues sur tout le texte sans redémarrage au début de chaque chapitre).
CO : <i>Arbre</i>		9 742 mots sur 61 531 mots du texte entier	Introduction du contexte du récit : CH1 : Rencontre imprévue des personnages principaux. CH2 : Changement de contexte spatio-temporel. Deuxième rencontre des personnages principaux et entrée en scène d'un personnage secondaire. CH3 : Changement de contexte spatio-temporel. Présentation des circonstances et des personnages secondaires. Troisième rencontre des personnages principaux.	

Tableau 19 Caractéristiques et contenus de notre corpus

Comme nous l'avons mentionné plus haut, ce corpus comprend deux textes inter-langues. Rappelons que la version originale du roman français est plus longue que celle du roman coréen. Proportionnellement parlant, le premier occupe un neuvième du roman entier alors que le deuxième en occupe un sixième ce qui correspond respectivement au début de la partie « perturbation » et à celle de « situation Initiale »³⁹.

3.3 Caractéristiques typiques des chaînes de référence dans le texte narratif

Capin et *al.* (2021) soulignent quelques traits typiques de ce domaine discursif : un nombre important de singletons, peu de chaînes longues et l'importance des pronoms personnels, catégorie grammaticale la plus représentée. Nous avons déjà examiné l'emploi du pronom personnel dans la section 2 *supra* et avons souligné que celui-ci n'a pas autant d'importance en coréen qu'en français. Dans ce qui suit, nous expliciterons donc deux autres caractéristiques du domaine discursif mentionnées par ces auteurs et la manière dont elles se manifestent réellement dans nos textes.

3.3.1 Le nombre de singletons

D'après la recherche menée sur les premières mentions par Manuélian et Schnedecker (2021) dont le corpus comprend quatre genres différents⁴⁰, 64,9 % des référents du début de texte restent à l'état de singletons dans les romans et seuls 34,1 % arrivent à constituer des chaînes. L'étude de Capin et *al.* (2021) sur un ensemble de textes narratifs du 12^{ème} au 20^{ème} siècles montre cette même tendance. La proportion de singletons se situe entre 71 % et 92 % selon l'époque. Dans cette recherche, les textes du 12^{ème} et du 15^{ème} siècles ont plus de singletons (environ 92% dans les deux cas) que les textes écrits à partir du 16^{ème} siècle.

Notre travail, rappelons-le, se concentre uniquement sur les chaînes, nous n'annotons ni les singletons ni les simples coréférences⁴¹. Pour ce faire, au début de l'annotation, nous avons marqué toutes les expressions référentielles et, à la relecture, nous avons enlevé celles

³⁹ Voir les points 1.3.3 et 1.3.4 du chapitre 5.

⁴⁰ Les auteurs travaillent sur le sous-corpus du projet DEMOCRAT, huit textes du 20^{ème} siècle de quatre genres différents (roman, presse, texte scientifique et texte juridique). Parmi les 960 premières mentions de tout le corpus, 305 sont extraites de romans.

⁴¹ Voir la section 1.2 du chapitre 6.

qui n'arrivaient pas à constituer de chaîne. La figure *infra* représente la première étape de l'annotation avant la relecture où nous avons annoté toutes les expressions référentielles. Il s'agit d'un des paragraphes du début du texte :

[#6] LogeSorelli La loge de pSorelli la Sorelli était d'Elegance une élégance officielle et banale . Psyche Une psyché , Divan un divan , Toilette une toilette et Armoires des armoires LogeSorelli en formaient MobilierNecessaire le mobilier nécessaire . Gravures Quelques gravures sur Murs les murs , souvenirs de pMere_Sorelli la mère , pMere_Sorelli qui avait connu BeauxJours_AncienOpera les beaux jours de AncienOpera l'ancien Opéra de RueLePeletier la rue Le Peletier . PortraitVestris Des portraits de Vestris , PortraitGardel de Gardel , PortraitDupont de Dupont , PortraitBigottini de Bigottini . LogeSorelli Cette loge paraissait Palais un palais Gamines_CDB aux gamines CorpsDeBallet du corps de ballet , Gamines_CDB qui étaient logées dans ChambresCommunes des chambres communes , ChambresCommunes où Gamines_CDB elles passaient Gamines_CDB leur temps à chanter, à se disputer, à battre Coiffeurs les coiffeurs et Habilleuses les habilleuses et à se payer PetitVerreCassis des petits verres de cassis ou PetitVerreBiere de bière ou même PetitVerreRhum de rhum jusqu' CoupDeCloche au coup de cloche de Avertisseur l'avertisseur .

Figure 30 Exemple d'annotation : un bon nombre de singletons (*Fantôme*)

Le nombre des singletons dans les cases grises⁴² (25 sur 36, 69,4 %) démontre que beaucoup de singletons se présentent dans les textes narratifs, comme le soulignent les auteurs précités. De nombreux référents (objets, circonstances, épisodes, etc.) participent à la construction du récit, plus précisément à la mise en place du contexte, et un grand nombre d'entre eux entrent en scène et disparaissent aussitôt selon leur rôle dans le récit. Les singletons tels « une psyché », « un divan », « une toilette », etc. sont mentionnés afin de fournir l'information sur le décor de la scène en cours⁴³.

3.3.2 La longueur des chaînes

Plus la chaîne est longue, plus le référent est mentionné. Cela signifie qu'il joue un rôle important dans le récit. Corollairement, dans les textes narratifs, les chaînes des personnages principaux sont les plus longues. Nous utilisons cette mesure pour trier les chaînes les plus

⁴² Il s'agit d'une partie extraite du corpus. L'expression « du corps de ballet » à la sixième ligne qui n'est apparue qu'une fois dans cet exemple n'est donc pas un singleton, parce qu'elle est coréférentielle avec plus de deux autres maillons extérieurs à cet exemple et constitue la chaîne « CorpsDeBallet » tout au long du texte.

⁴³ Dans la recherche de Manuélian et Schnedecker (2021 : 225), environ 20 % des entités humaines, des objets concrets et des lieux sont des singletons. La proportion de singletons est plus élevée (30 %) dans les cas de référents abstraits comme des moments ou des idéalités.

importantes, notamment, celles des personnages principaux. Tenant compte du volume de nos textes, nous définissons les chaînes courtes comme celles qui comprennent trois ou quatre maillons et les chaînes longues comme celles qui ont plus de 30 maillons⁴⁴. Le tableau suivant détaille le nombre de chaînes selon la longueur :

Nombre de chaînes		Chaînes courtes : 3 ou 4 maillons	Chaînes moyennes : de 5 à 29 maillons	Chaînes longues : plus de 30 maillons
<i>Fantôme</i>	162	77 (47,5 %)	72 (44,4 %)	13 (8 %)
<i>Arbre</i>	185	116 (62,7 %)	63 (34 %)	6 (3,2 %)

Tableau 20 Nombre de chaînes selon la longueur

D'après les pourcentages de chaînes longues représentés dans le tableau 20, nous observons respectivement 8 % des chaînes longues dans le roman français et 3,2 % dans le coréen, ce qui justifie le propos de Capin et *al.* (2021) sur les traits typiques des textes narratifs.

Parmi 162 chaînes détectées en français, les 13 chaînes représentées dans le tableau qui suit correspondent aux référents les plus souvent mentionnés.

Type de chaînes		Nom des chaînes	Nombre de maillons	Distance inter-maillonnaire
Contextes spatio-temporels		Opéra	30	356 mots
Personnages	Individu	Sorelli	100	89 mots
		Petite Jammes	40	231 mots
		Fantôme	170	60 mots
		Petite Giry	44	196 mots
		Joseph Buquet	51	190 mots
		Gabriel	36	12 mots
		Moncharmin	30	236 mots

⁴⁴ Par chaîne longue, Capin et *al.* (2021) entendent les chaînes de « plus de 10 maillons ». Pourtant, nous avons pensé que ce critère ne peut pas indiquer un résultat significatif dans notre cas où la taille moyenne des chaînes est de 12,3 maillons dans le texte français et de 10,6 maillons dans le texte coréen. Dans *Fantôme*, les 39 chaînes ayant plus de 10 maillons sont celles des personnages (principaux et secondaires), des contextes spatio-temporels et des motifs (objets, incidents, etc.). Pourtant, tous les référents n'ont pas la même importance dans les romans. Dans des textes relativement longs comme les nôtres, le critère « plus de 10 maillons » ne suffit pas à trier les chaînes les plus importantes, qui sont, en règle générale, celles des personnages. Considérant la longueur du corpus (environ 10 000 mots) et la longueur moyenne des chaînes (12,3 et 10,6 maillons), nous avons décidé de trier les chaînes longues de plus de 30 maillons, sans quoi notre intention de trier les chaînes les plus importantes et d'exclure celles qui sont secondaires ne sera pas bien réalisée.

		Comte de Chagny	85	36 mots
		Christine Daaé	128	33 mots
		Raoul	223	14 mots
	Groupe	Ex-directeurs	81	130 mots
		Nouveaux directeurs	33	215 mots

Tableau 21 Les CR de plus de 30 maillons (*Fantôme*)

Dans ce tableau, seule la chaîne qui renvoie à *Opéra*, lieu du récit, entre dans la catégorie de contexte spatio-temporel ; les autres désignent les personnages. Les trois chaînes les plus longues dans le texte *Fantôme*, *fantôme*, *Christine Daaé* et *Raoul*, désignent, en effet, les personnages principaux du roman. En outre, leur distance inter-maillonnaire relativement courte montre également que ce sont les chaînes les plus importantes dans lesquelles les maillons se présentent, pour parler du cas de *Fantôme*, tous les 60 mots (approximativement toutes les trois lignes).

Type de chaînes		Nom des chaînes	Nombre de maillons	Distance inter-maillonnaire
Personnages	Individu	Chef de Journal X	30	67 mots
		Dong-hyuk	467	19 mots
		Young-shin	280	28 mots
		Mme Baek	86	41 mots
		Frère de DH	46	54 mots
	Groupe	YS et DH	38	206 mots

Tableau 22 Les CR de plus de 30 maillons (*Arbre*)

Dans « Arbre », seules les chaînes renvoyant à des personnages ont plus de 30 maillons, ce qui semble normal, d'après ce que nous avons mentionné dans le tableau 19 *supra*. En effet, les contextes spatio-temporels des trois premiers chapitres du roman ne sont pas identiques, autrement dit, dans chaque chapitre et même à l'intérieur d'un même chapitre, le récit se déroule à différentes dates et dans différents lieux. La portée des chaînes, est donc, dans de nombreux cas, locale. Il en va de même pour les chaînes relatives aux motifs liés aux contextes comme le décor de la maison ou la tenue d'un personnage, elles sont généralement courtes.

3.4 Le contexte spatio-temporel et les chaînes de référence

La structuration discursive joue un rôle important dans la compréhension de la nature des

CR. Elle concerne non seulement la façon dont les marqueurs référentiels sont distribués (Schneedecker, 1997 ; Baumer 2015, voir la section 2.5 *supra*) mais aussi fournit une explication sur certains paramètres. D'après nos observations, le changement du contexte spatio-temporel qui fonctionne comme cadre des événements narrés peut influencer certaines mesures telles que la longueur et la portée des CR. L'exemple (33) montre quelques changements dans le contexte spatio-temporel du premier chapitre du texte *Arbre* :

(33) [*Arbre*, #3, #136, #219 et #234]

[#3] 가을 학기가 되자, ○○일보사에서 주최하는 학생계몽운동에 참가하였던 대원들이 돌아왔다. **오늘 저녁 (aujourd'hui soir)**은 각처에서 모여든 대원들을 위로하는 다과회가 **그 신문사 누상 (bâtiment de ce journal)**에서 열린 것이다.

À la rentrée scolaire, les militants qui avaient participé au mouvement de l'éclaircissement organisé par le Journal OO sont rentrés. **Ce soir**, une réunion pour encourager ces militants venant de partout a lieu dans **le bâtiment de ce journal**.

[#136] 매우 긴장된 중에 K보육학교 학생들의 코러스로 간친회는 파하였다. 동혁은 여러 학생들 틈에 섞여서 **서대문행 전차 (tram direction Seodaemun)**를 탔다.

Dans une ambiance tendue, la réunion se termine par une représentation donnée par le chœur du Lycée K de l'éducation des enfants. Dong-hyuk est monté dans **le tram direction Seodaemun** avec d'autres étudiants.

[#219] **그 이튿날 (lendemain)**, **학교로 (à lycée)** 내려간 뒤에, 동혁은 며칠 동안 마음의 안정을 잃고 지냈다.

Le lendemain, après être retourné **au lycée**, Dong-hyuk a perdu (son) calme intérieur pendant quelques jours.

[#234] 그러다가, **하루 (un jour)**는 천만뜻밖에 영신에게서 편지가 왔다.

Un jour, une lettre inattendue est arrivée de Young-shin.

Les paragraphes de [#3] à [#218] décrivent des incidents qui ont lieu durant une même soirée, le soir de la réunion. Mais, à partir du paragraphe [#219], les événements se passent le lendemain de cette réunion. Au paragraphe [#234], « Un jour » indique une modification dans la chronologie du récit.

Par ailleurs, le récit se déroule dans plusieurs lieux différents. Du paragraphe [#3] jusqu'au [#136], le contexte spatial est le bâtiment du journal où se passe la réunion. Mais les

changements se manifestent au [#136] et au [#219] : « le tram direction Seodaemun » et « le lycée de Dong-hyuk ».

Le changement fréquent du contexte spatio-temporel peut influencer la longueur et la portée d'une chaîne. Par exemple, les personnages ou les objets mentionnés dans la scène de la réunion (« le chef du journal » ou « l'estrade de l'amphi ») disparaissent à partir du paragraphe [#136] où le contexte spatial change et se centre sur le « tram direction Seodaemun ». Ainsi, les chaînes fortement dépendantes d'un certain contexte spatio-temporel se ferment lors d'un changement de scène.

Ainsi, et comme vu plus haut dans le tableau 20, nous comptons davantage de chaînes courtes dans *Arbre* que dans *Fantôme* : 62,7% dans le premier et 47,5% dans le deuxième. Le texte coréen, au regard du texte français, se caractérise par moins de mots, plus de chaînes et moins de chaînes longues (ou bien plus de chaînes courtes). Comme le soulignent Capin *et al.* (2021 : 95), la présence de chaînes courtes est étroitement liée à « la présence épisodique de certains personnages ou objets ». Ce caractère découle du fait que *Arbre* a des motifs plus variés. En effet, le roman consiste davantage en sous-thèmes, épisodes et circonstances, comme nous venons de le voir dans l'exemple (33), qui est un ensemble de paragraphes extraits uniquement du chapitre 1 du roman. Cette tendance continue dans les chapitres 2 et 3. Le tableau 23 *infra* récapitule les éléments constituant nos romans :

	<i>Fantôme</i>	<i>Arbre</i>
Personnages principaux	Trois personnages principaux (Fantôme, Raoul et Christine)	Deux personnages principaux (Dong-hyuk et Young-shin)
Date (Durée)	Une soirée (pendant le gala)	Approximativement dix mois (Le récit commence à la fin de l'été et continue jusqu'au début de l'été de l'année suivante)
Lieux (Rayon d'action des personnages)	Ch1 : 1 endroit (loge de la Sorelli) Ch2 : 4 endroits dans l'Opéra (foyer de la danse, scène, loge de Daaé et couloir) Ch3 : 2 endroits (foyer de la danse et bureau directorial)	Ch1 : 4 endroits (salle de réunion, tram, rue et ville de Dong-hyuk) Ch2 : 3 endroits à Séoul (chez Mme Baek, stade et vallée Acbak) Ch3 : 4 endroits dans la ville de Dong-hyuk (chez Dong-hyuk, plage, port et chambre de Young-shin)
Incident (Episode)	Ch1 : 1 incident (mort de Joseph Buquet)	Ch1 : 3 incidents (réunion, rencontre dans le tram et dialogue dans la rue)

	Ch2 : 3 incidents (succès de Daaé, rencontre de Daaé & Raoul et déception de Raoul) Ch3 : 2 incidents (apparition du fantôme et dialogue des directeurs)	Ch2 : 3 incidents (match de foot, dîner chez Mme Baek et serment de Dong-hyuk & Young-shin) Ch3 : 2 incidents (pluie et visite de Young-shin)
--	---	--

Tableau 23 Les éléments susceptibles d'influencer la longueur des chaînes

Par cette analyse, nous pouvons, avec prudence, présumer que *Arbre* a une vitesse de narration supérieure à celle de *Fantôme*. Dans *Arbre*, le temps du récit est plus rapide ; à peu près 10 mois se passent en trois chapitres alors que tous les événements se passent en une soirée dans *Fantôme*.

En outre, davantage d'endroits différents⁴⁵ sont introduits dans le texte *Arbre*. Le changement de scène est réalisé par le changement de lieu et ces endroits arrivent à constituer une chaîne. Cet aspect mérite d'être souligné, car, lorsque de nouvelles circonstances sont données, notamment dans des romans où la description autour de personnages et d'incidents prend une place importante, le narrateur doit faire plus d'efforts pour la mise en place du contexte. L'exemple (34) *infra* montre cet effort :

(34) [*Arbre*, #295 et #296]

[#295] …동혁은 두리번거리며 [_{chez MmeB} 집(maison)] 안을 둘러본다. [_{chez MmeB} 삼천 원이나 들여서 새로 지었다는 집 (maison dont le prix de la reconstruction coûte trois milles won)]은 1네 귀(quatre angles d'arêtier)가 반짝 들렸는데, 2서까래 (chevonnages)까지 비둘기장처럼 파란 뽕끼칠을 하였고, 3[Salon MmeB 분합 마루(salle ouverte)] 유리창(vitraux de la salle ouverte)에는 장미꽃 무늬가 혼란한 휘장을 늘어 쳤다. 4마당(jardin)은 그다지 넓지 못하나 각색 화초가 어울려 피었는데, 그 중에도 이름과 같이 청초한 옥잠화 두어 분은 황혼에 그윽한 향기를 놓는다.

Dong-hyuk jette des regards de tous les côtés de [_{chez MmeB} la maison]. [_{chez MmeB} La maison dont le prix de la reconstruction coûte trois milles won] a 1les quatre angles d'arêtier (du toit) levants vers le ciel et est peinte en bleu même jusqu'2aux

⁴⁵ Le terme « endroit » ici, ne signifie pas tous les lieux mentionnés (soit par le narrateur soit par les personnages) dans le récit. Il réfère aux lieux où se situent les personnages et où se passent les incidents, en d'autres termes, le rayon d'action des personnages.

chevronnages comme si on était dans la cage des colombes. ³Les vitraux de [Salon MmeB **la salle ouverte**] ont de splendides dessins de roses. ⁴Le jardin n'est pas grand mais diverses fleurs sont plantées et parmi ces fleurs, deux Hosta, purs, comme le signifie leur nom, ajoutent une odeur profonde au crépuscule.

[#296] 먼저 온 회원들은 [Salon MmeB **응접실로 쓰는 대청(grande pièce qui sert de salon)**]에 모여서 혹은 [Piano **피아노(piano)**]를 눌러 보고, 혹은 백씨가 구미 각국으로 시찰과 강연을 하러 다닐 때 박힌 사진첩을 꺼내 놓고 둘러앉았다.

Les premiers arrivés réunis dans [Salon MmeB **la grande pièce qui sert de salon**] effleurent les touches [Piano **du piano**] ou regardent les photos que Mme Baek a prises à l'étranger.

Dans cet exemple, le cadre spatial du récit est « la maison de Mme Baek ». Pour l'introduire, d'autres chaînes courtes associées peuvent être créées : les chaînes « salon Mme Baek » (écrit en rouge) et « piano » (en bleu). Le nombre de maillons de chaque chaîne est respectivement de cinq et trois. À part ces chaînes, les singletons renvoyant aux référents associés à la maison sont observés : « les quatre angles d'arêtier », « les chevronnages », « les vitraux de la salle ouverte⁴⁶ » et « le jardin ».

Ainsi, les données sur la longueur des chaînes et leur portée aident à comprendre non seulement les différents degrés d'importance des référents participant au déroulement du récit, mais aussi le découpage épisodique ou la structure narrative de celui-ci.

3.5 Les chaînes de référence des personnages dans les romans

Comme nous venons de le remarquer avec les tableaux 21 et 22 plus haut, dans les textes narratifs, les chaînes les plus longues sont souvent celles des personnages principaux. Dans cette partie, nous allons préciser les caractéristiques générales de ces chaînes et leurs particularités.

3.5.1 Modalités d'introduction des personnages

D'après le classement des expressions référentielles de Charolles (1987), certaines

⁴⁶ La salle ouverte est un espace appartenant à la maison traditionnelle en Corée qui joue le rôle de salon.

d'entre elles, les SN indéfinis, les noms propres, les descriptions définies complètes et les expressions nominales démonstratives (usage déictique ou ostensif), sont « susceptibles d'apparaître en première ou en deuxième et énième mention d'une chaîne de référence » (Schneedecker, 2021 : 80). Par ailleurs, les noms propres, les descriptions définies complètes, les descriptions définies incomplètes, les expressions nominales démonstratives, les descriptions indéfinies, les noms de qualité et les pronoms sont des expressions qui peuvent se placer comme deuxième et énième maillon. Le tableau suivant proposé par Schneedecker (2021 : 82) récapitule les expressions référentielles qui peuvent servir de 1^e ou 2/n^{ième} maillon :

1 ^e maillon	2/n ^{ième} maillon
Nom propre	
Descriptions définies complètes	
	Descriptions définies incomplètes
SN démonstratifs (usage déictique)	SN démonstratifs
SN indéfinis	
	SN dits de qualité
	Pronoms

Tableau 24 Expressions référentielles et leur position possible dans la chaîne (Schneedecker, 2021)

Pourtant, l'auteur insiste sur certaines exceptions découlant de raisons stylistiques, notamment dans les romans. Dans ce qui suit, nous examinerons la distribution des premiers maillons des chaînes des personnages dans nos textes.

3.5.1.1 Les maillons initiaux des chaînes des personnages

Comme le souligne Schneedecker (2021), dans les romans, les maillons initiaux peuvent varier selon le propos du locuteur. Dans l'exemple (35) ci-après, le pronom personnel *je* est employé en tant que premier maillon désignant le personnage principal dans la première phrase du roman :

(35) Puis-**je**, monsieur, vous proposer **mes** services, sans risquer d'être importun ? — (la première phrase du roman « La chute » par Albert Camus).

Les SN indéfinis qui ont normalement pour objectif d'introduire un nouveau référent, peuvent se présenter au milieu d'une chaîne en tant qu'énième maillon. C'est bien le cas du deuxième « un lieutenant de pompiers » de (36) :

(36) [*Fantôme*, #24 et #25]

Un lieutenant de pompiers, c'est brave ! Ça ne craint rien, ça ne craint surtout pas le feu !

Eh bien, le lieutenant de pompiers en question, qui s'en était allé faire un tour de surveillance dans les dessous et qui s'était aventuré, paraît-il, un peu plus loin que de coutume, était soudain réapparu sur le plateau, pâle, effaré, tremblant, les yeux hors des orbites, et s'était quasi évanoui dans les bras de la noble mère de la petite Jammes. Et pourquoi ? Parce qu'il avait vu s'avancer vers lui, à hauteur de tête, mais sans corps, une tête de feu ! Et je le répète, **un lieutenant de pompiers, ça ne craint pas le feu.**

Dans le tableau 25 qui suit, nous énumérons le premier et le deuxième maillon des chaînes des personnages de nos textes. Considérant le fait que ceux-ci ne comportent respectivement que trois et deux personnages principaux, nous avons élargi le champ des exemples afin de voir la tendance générale de la première occurrence des personnages dans le roman. L'examen porte sur les chaînes des huit personnages de chaque texte les plus souvent mentionnés.

Texte	Personnage	1 ^e maillon	2 ^e maillon
<i>Fantôme</i>	Raoul	Son frère Raoul	Raoul
	fantôme	Ce (c'est le fantôme !)	Le fantôme
	Christine Daaé	Christine Daaé	Elle
	Sorelli	La Sorelli	La Sorelli
	Comte de Chagny	Le comte de Chagny qui montait	Le comte, ordinairement si calme
	Joseph Buquet	Joseph Buquet, chef machiniste, qui...	Il
	Petite Giry	La petite Giry	Sa
	Petite Jammes	La petite Jammes	Qui
<i>Arbre</i>	Dong-hyuk	(la) Personne qui a eu le meilleur résultat parmi les militants de cette année et qui a aussi établi un record ...	Dong-hyuk Park du lycée XX.
	Young-shin	Young-shin Chae du lycée chrétien des filles qui a eu le meilleur résultat...	(la) Fille appelée 'Young-shin Chae'

Mme Baek	Mme Baek	Mme Baek
Dong-hwa (frère de Dong-hyuk)	Dong-hwa	∅ (pronom zéro)
Chef de Journal X	(le) Chef de ce Journal	∅ (pronom zéro)
Kun-bae	Kun-Bae	(un) Collègue qui s'appelle Kun-Bae
Monsieur Park	Monsieur Park qui était en train de creuser le champ d'orge couvert par les poussières de sable sec	Il
Madame Park	(sa) Femme qui vient d'entrer avec des algues	Il (équivalent à « elle »)

Tableau 25 Le premier et le deuxième maillon de chaque chaîne sur le personnage

Les expressions dans les cases grises sont celles qui ne sont pas parfaitement adaptées au classement de Charolles (1987). En suivant le point de vue de Schnedecker (2021), nous les considérons comme des exceptions dues à une raison stylistique.

Parmi les expressions employées en tant que premier maillon, nous rencontrons souvent des formes particulières. Examinons tout d'abord la chaîne de *Raoul* dans le texte *Fantôme* :

(37) [*Fantôme*, #120 et #121]

La fortune des Chagny était considérable, et quand le vieux comte, qui était veuf, mourut, ce ne fut point une mince besogne pour Philippe, que celle qu'il dut accepter de gérer un aussi lourd patrimoine. Ses deux sœurs et **son frère Raoul** ne voulurent point entendre parler de partage, et ils restèrent dans l'indivision, s'en remettant de tout à Philippe, comme si le droit d'aînesse n'avait point cessé d'exister. Quand les deux sœurs se marièrent, – le même jour, – elles reprirent leurs parts des mains de leur frère, non point comme une chose leur appartenant, mais comme une dot dont elles lui exprimèrent leur reconnaissance.

La comtesse de Chagny – née de Moerogis de la Martynière – était morte en donnant le jour à **Raoul**, né vingt ans après son frère aîné. Quand le vieux comte était mort, **Raoul** avait douze ans.

Dans l'exemple (37), le narrateur décrit l'histoire de la famille Chagny, notamment celle de *Philippe*, et au sein de cette description, *Raoul* est présenté pour la première fois en tant

que petit frère de *Philippe*, « son frère Raoul ». Le regard du narrateur passe ainsi de *Philippe* à *Raoul*. Pour faire le pont entre l'histoire précédente et celle de Raoul, le narrateur choisit alors la forme de coordination « SN + NPR ». Celle-ci justifie la relation entre les deux référents. Une fois introduit dans le texte par cette mention, le référent est repris par le nom propre « Raoul » dans les deuxième et troisième maillons.

Dans l'exemple suivant, le personnage *Joseph Buquet* est mentionné, pour la première fois, sous la forme du sujet inversé introduit par un Npr suivi d'une apposition⁴⁷.

(38) [*Fantôme*, #18]

Tout cela était-il sérieux ? La vérité est que l'imagination du squelette était née de la description qu'avait faite du fantôme **Joseph Buquet, chef machiniste, qui, lui, l'avait réellement vu. Il s'était heurté**, – on ne saurait dire « nez à nez », car le fantôme n'en avait pas, – avec le mystérieux personnage dans le petit escalier qui, près de la rampe, descend directement au « dessous ».

Le narrateur, en introduisant le personnage *Joseph Buquet* dans le récit, transmet l'information complémentaire *via* l'insertion d'une apposition, sinon l'occurrence simple de « Joseph Buquet » semblerait importune, puisque l'on ne dispose d'aucune connaissance préalable sur ce personnage.

En coréen, certaines chaînes débutent par des constructions appositives. L'exemple (39) *infra* est le début du troisième chapitre du roman *Arbre* qui décrit le nouveau contexte spatio-temporel et les nouveaux personnages secondaires autour du personnage principal, *Dong-hyuk* :

(39) [*Arbre*, #452, #453 et #454]

"허어 이 날, 사람을 잡으려구 이렇게 가무는 게여." 바싹 마른 흙이 먼지처럼 피어올라 폴삭폴삭 날리는 보리밭에 붓을 주던 박첨지 (**Monsieur Park qui était en train de creuser le champ d'orge couvert par les poussières de sable sec**)는, 기신없이 괜이질을 하던 손을 쉬고 허리를 펴며 (Ø) 혼자말로 탄식을 한다. 그(ii)는 검버섯이 돋은 이마에 주름살을 잡으며 머리 위를 우러러본다. 그러나 가을날처럼

⁴⁷ Voir la section 2.2.6.2 du chapitre 6.

새파란 하늘에는 구름 한 점 찾아 낼 수가 없다. 바닷가의 메마른 농촌에 바람만 진종일 씩씩 불어서 콧구멍이 막히고 목의 침이 말라드는 것 같다.

« Hélas, cette sécheresse tuera l'homme ». **Monsieur Park qui était en train de creuser le champ d'orge couvert par les poussières de sable sec**, arrête (sa) houe à main en se redressant et (∅) soupire. Il regarde au-dessus de (sa) tête en mettant la main sur la ride de (son) front où il y a la tâche sénile. Pourtant, dans le ciel bleu comme celui d'automne, on ne trouve même pas un morceau de nuage. Dans le village rural près de la mer, il n'y a que le vent sec qui fait qu'on a l'impression d'avoir le nez bouché et la gorge séchée.

"이런 제에기, 보리싹이 연골에 말려 배틀어지니 올 여름엔 냉수만 마시구 산담메." 늙은이(vieux)는 다시 한번 말과 한숨을 뒤섞어 내뿜고는, 이제야 겨우 강아지풀 앞사귀만하게 꼬리를 흔드는 보리싹을 짚신발로 걷어찬다.

« Zut, l'herbe d'orge est toute séchée. On n'a que de l'eau à boire cet été ». **Le vieux** soupire encore une fois en parlant et (il) donne un coup de pied aux herbes qui ont à peu près la même taille que la feuille de vulpin.

La chaîne du nouveau personnage *Monsieur Park* commence après le monologue. Son premier maillon est un nom propre expansé par une relative décrivant l'action du personnage et les circonstances : « Monsieur Park qui était en train de creuser le champ d'orge... ». Ensuite, le référent est repris par le pronom zéro et le pronom personnel.

D'après nos observations, il n'y a pas de grande différence à noter au niveau formel des maillons initiaux dans les deux textes. Cependant, nous croyons qu'il est essentiel de considérer les premiers maillons dans les romans en rapport avec leurs contextes d'occurrence, en d'autres termes, il faudrait discerner le cas où le personnage est introduit par le discours d'un autre personnage de celui où il est décrit par le narrateur.

3.5.1.2 Le niveau énonciatif d'apparition des personnages

Le tableau 26 suivant propose un classement des premiers maillons des personnages mentionnés dans le tableau 25 *supra* selon leurs locuteurs :

	1 ^{ère} mention introduite par un personnage	1 ^{ère} mention introduite par le narrateur
<i>Fantôme</i>	<ul style="list-style-type: none"> - fantôme : Ce - Christine Daaé : Christine Daaé 	<ul style="list-style-type: none"> - Raoul : Son frère Raoul - Sorelli : La Sorelli, un des premiers sujets de la danse... - Comte de Chagny : Le comte de Chagny qui montait - Joseph Buquet : Joseph Buquet, chef machiniste... - Petite Giry : La petite Giry, – des yeux pruneaux, des cheveux d'encre, un teint de bistre, ... - Petite Jammes : La petite Jammes, – le nez cher à Grévin, des yeux de myosotis, des joues de roses, ...
<i>Arbre</i>	<ul style="list-style-type: none"> - Dong-hyuk : <u>(la) Personne qui a eu le meilleur résultat parmi les militants ...</u> - Young-shin : <u>Young-shin Chae du lycée chrétien des filles qui a eu le meilleur résultat...</u> - Mme Baek : Madame Baek - Dong-hwa: Dong-hwa 	<ul style="list-style-type: none"> - Chef de Journal X : (le) Chef de ce Journal - M Park : Monsieur Park qui était en train de creuser le champ d'orge couvert par les poussières de sable sec - Mme Park : (sa) Femme qui vient d'entrer avec des algues - Kun-bae : Kun-bae

Tableau 26 Différenciation des premières mentions selon le niveau énonciatif d'apparition

Dans *Fantôme*, la plupart des personnages sont introduits et décrits au niveau de la narration sauf le *fantôme* et *Christine Daaé*. Regardons les deux exemples suivants :

(40) [*Fantôme*, de #4 à #7]

Ce fut la petite Jammes, – le nez cher à Grévin, des yeux de myosotis, des joues de roses, une gorge de lis, – qui en donna la raison en trois mots, d'une voix tremblante qu'étouffait l'angoisse : « **C'est le fantôme !** »

Et elle ferma la porte à clef.

[...]

La Sorelli était très superstitieuse. En entendant la petite Jammes parler **du fantôme**, elle frissonna et dit : « Petite bête ! »

Et comme elle était la première à croire aux fantômes en général et à **celui de l'Opéra** en particulier, elle voulut tout de suite être renseignée.

(41) [*Fantôme*, #105, #106 et #107]

Au premier palier, la Sorelli se heurta au comte de Chagny qui montait. Le comte, ordinairement si calme, montrait une grande exaltation.

« J'allais chez vous, fit le comte en saluant la jeune femme de façon fort galante. Ah ! Sorelli, quelle belle soirée ! Et **Christine Daaé** : quel triomphe !

– Pas possible ! protesta Meg Giry. Il y a six mois, **elle** chantait comme un clou !

Comme nous le voyons dans les exemples (40) et (41), les deux personnages principaux du roman, *fantôme* et *Christine Daaé*, sont introduits *via* le discours rapporté d'un autre personnage. En l'occurrence, ils sont donnés de manière simple, c'est-à-dire par les expressions référentielles de forme relativement courte. Cela tient au fait que le dialogue des personnages présuppose une connaissance préalablement partagée entre eux.

En revanche, les lecteurs ne disposant pas de cette connaissance partagée, ont besoin d'une explication plus précise sur le référent en question. Le narrateur doit donc fournir les informations nécessaires sur ce référent. C'est pour cette raison que la première occurrence qu'il donne à propos d'un référent a souvent une forme moins simple que celle introduite par un personnage, notamment, dans des romans comme les nôtres marqués par le point de vue omniscient. Nous avons déjà remarqué que le premier maillon de la chaîne de *Raoul* de (37) *supra* se présente sous la forme de « SN + NPR ».

Dans les chaînes des personnages secondaires, comme celle de (38), le NPR accompagné d'une apposition est aussi fréquemment utilisé en tant que premier maillon. L'exemple (42) montre les deux premiers paragraphes du roman où l'on introduit quelques personnages secondaires :

(42) [*Fantôme*, #3 et #4]

Ce soir-là, qui était celui où [Ex-directeurs **MM. Debienne et Poligny**], **les directeurs démissionnaires de l'Opéra**, donnaient leur dernière soirée de gala, à l'occasion de

leur départ, la loge de [La Sorelli **la Sorelli**], *un des premiers sujets de la danse*, était subitement envahie par une demi-douzaine de ces demoiselles du corps de ballet qui remontaient de scène après avoir « dansé » Polyeucte. Elles s’y précipitèrent dans une grande confusion, les unes faisant entendre des rires excessifs et peu naturels, et les autres des cris de terreur.

La Sorelli, qui désirait être seule un instant pour « repasser » le compliment qu’elle devait prononcer tout à l’heure au foyer devant MM. Debienne et Poligny, avait vu avec méchante humeur toute cette foule étourdie se ruer derrière elle. Elle se retourna vers ses camarades et s’inquiéta d’un aussi tumultueux émoi. Ce fut [La Petite Jammes **la petite Jammes**], – *le nez cher à Grévin, des yeux de myosotis, des joues de roses, une gorge de lis*, – qui en donna la raison en trois mots, d’une voix tremblante qu’étouffait l’angoisse : « C’est le fantôme ! »

Au cours de ces paragraphes, trois personnages (ou groupe de personnages) sont introduits pour la première fois par le narrateur. Afin d’aider les lecteurs à se plonger immédiatement dans le récit, il choisit de présenter le nom du personnage avec une description essentielle de cette personne ou de contexte dans lequel elle se présente.

En revanche, les chaînes des personnages principaux de texte *Arbre* montrent une disparité : l’occurrence du premier maillon des chaînes *Dong-hyuk* et *Young-shin*, est effectuée *via* le discours rapporté. Cependant, contrairement à ce que nous venons de voir avec la chaîne de *fantôme* ou celle de *Christine Daaé*, les premières mentions de ces chaînes ont respectivement la forme « Relative + N » et « Relative + Npr » fonctionnellement identiques à la construction appositive en français.

(43) [*Arbre*, #30 et #31]

"금년에 활동한 계몽 대원 중에 뛰어나게 좋은 성적을 보여 주었을 뿐 아니라, 글을 깨쳐 준 아동의 수효로는 우리 신문사에서 이 운동을 개시한 이래 최고 기록을 지은 분 (**personne qui a non seulement eu le meilleur résultat parmi les militants de cette année mais aussi a établi un record en nombre d’enfants alphabétisés depuis que notre Journal a commencé ce mouvement**)을 (Ø) 소개 하겠소이다."

“(Je) (vous) introduis **la personne qui a non seulement eu le meilleur résultat parmi les militants de cette année mais aussi a établi un record en nombre**

d'enfants alphabétisés depuis que notre Journal a commencé ce mouvement."

하고는 (Ø) 다시 안경 너머로 서류를 들여다보다가 얼굴을 들고 선생이 출석부를 부르듯이, "××고등농림의 박동혁군 (**M Park de lycée agricole XX**)!" 하고 (Ø) 목소리를 높였다.

Après avoir dit ça, (il) a regardé par-dessus (ses) lunettes le document et (il) a haussé la voix en levant (sa) tête comme s'(il) était le professeur qui appelle un étudiant. « **M Park de lycée agricole XX** ».

(44) [*Arbre*, de #99 à #102]

사회자는, "이번에는 금년에 처음으로 참가한 여자 대원 중에서 제일 좋은 성적을 나타낸 ××여자신학교에 재학중인 채영신양(**Young-shin Chae de XX lycée chrétien des filles qui a eu le meilleur résultat parmi les militantes ayant participé pour la première fois**) 의 감상담이 있겠습니다." 하고 오른편에 여자들이 모여 앉은 데를 (Ø) 바라다본다.

Le modérateur a dit, « Cette fois-ci, c'est la présentation de **Young-shin Chae de XX lycée chrétien des filles qui a eu le meilleur résultat parmi les militantes ayant participé à ce mouvement pour la première fois** » et (il) regarde le coin où les filles sont assises.

남학생들은 그편으로 머리를 돌리며 손뼉을 친다. '채영신'이라고 불린 여자 (**fillette nommée 'Young-shin Chae'**)는 한참 만에 얼굴이 딸깃빛이 되어 가지고 일어나더니, "전 아무 말도 하기 싫습니다!" 하고 (Ø) 머리를 내저으며 여무지게 한마디를 하고는 (Ø) 펄쩍 주저앉아 버린다.

Les garçons applaudissent en tournant leurs têtes vers la direction. **La fille nommée 'Young-shin Chae'** s'est levée après un moment avec une tête rouge, (elle) a dit avec rigueur en secouant la tête, « **Je** ne veux rien dire ! » et (elle) s'est assise.

Dans les deux cas, la première mention se présente au sein du discours d'un personnage, *le chef du Journal* qui joue le rôle du modérateur de la réunion. Ce contexte d'occurrence explique la raison de la forme complexe de la première mention introduite par un autre personnage du roman et donc qui, par hypothèse, et en fonction de ce que nous avons démontré auparavant, pourrait normalement être simple. Nous avons souligné plus haut que,

en règle générale, les personnages du roman sont considérés comme partageant certaines connaissances préalables sur le personnage principal, en d'autres termes, ils sont censés disposer de davantage de savoirs que les lecteurs qui se situent à l'extérieur du récit. Ce qui rend la première mention de chaque chaîne dans (43) et (44) exceptionnelle est le fait que *le chef du Journal*, qui mentionne *Dong-hyuk* et *Young-shin* pour la première fois, les introduit au public qui ne les connaît pas encore ; il joue un rôle similaire à celui du narrateur du roman.

Par ailleurs, l'exemple (45) montre le référent introduit par le discours d'un autre personnage : le personnage *Madame Baek* est introduit par *Young-shin* via le nom propre « Madame Baek ».

(45) [*Arbre*, de #185 à #192]

영신은, "저 때문에 너무 걱정을 허셔서 미안합니다." 하고는 구둣부리로 길바닥을 후비듯 하다가, 고개를 외로 꼬고 무엇을 생각하더니, "인전 **백선생님 (Mme Baek)** 집으로나 갈까 봐요." 한다.

Young-shin a dit en frottant la terre avec son talon, « Désolée de vous faire vous inquiéter pour moi », (elle) a réfléchi avec sa tête inclinée à gauche, et (elle) a dit « Maintenant (je) vais chez **Mme Baek** ».

"**백선생(Mme Baek)**이라니요?"

« **Mme Baek ?** »

"왜 여자기독교연합회 총무로 있는, **백현경 씨 (Madame Baek Hyun-kyung)**를 모르세요?"

« (Vous) ne connaissez pas **Madame Baek Hyun-kyung**, qui est le secrétaire général de l'Association des Chrétiennes ? »

"이름은 익숙히 들었지만…… **그(lui)**의 집이 이 근처가요?"

« (J)'ai entendu son nom... **Sa** maison est près d'ici ? »

Comme le montre l'exemple ci-dessus, lorsque le référent est donné pour la première fois dans le dialogue des personnages, le locuteur tend d'abord à le mentionner par le nom propre puis, en fonction du degré de connaissance de l'interlocuteur, à ajouter des informations supplémentaires sur ce référent au cours du dialogue. Reprenons (31) issu de *Fantôme* qui est un cas semblable à (45) :

(31) Au premier palier, la Sorelli se heurta au comte de Chagny qui montait. Le comte, ordinairement si calme, montrait une grande exaltation.

« J'allais chez vous, fit le comte en saluant la jeune femme de façon fort galante. Ah ! Sorelli, quelle belle soirée ! Et **Christine Daaé** : quel triomphe !

– Pas possible ! protesta Meg Giry. Il y a six mois, **elle** chantait comme un clou !

Ainsi, nous constatons donc que le propos de Schnedecker (2021) sur les types d'expressions référentielles pouvant servir de premier maillon de chaîne convient à nos textes. Les maillons initiaux peuvent être variés dans les romans. Nous avons également remarqué que les premiers maillons des CR des personnages, entre autres, se présentent différemment selon leurs contextes d'occurrence. À ce propos, deux cas sont testés : le personnage introduit par le discours d'un autre personnage et celui présenté par le narrateur. D'après nos observations *via* le tableau 26 et les deux exemples de (43) et (44), il est difficile de généraliser le contexte d'apparition des personnages principaux et celui des personnages secondaires. Cependant, les cas testés montrent une nette disparité : un personnage se présente sous une forme relativement simple lorsqu'il est introduit par le dialogue d'un autre, mais quand le référent est donné par le narrateur (ou le personnage qui joue ce rôle dans le récit), la forme d'introduction est souvent plus longue (« Relative + N/Npr »). De ce fait, nous pouvons souligner l'importance du niveau énonciatif dans le choix de la description utilisée pour introduire un personnage pour la première fois dans nos textes.

3.5.2 Les anaphores nominales des chaînes des personnages

Reprenons maintenant les chaînes des personnages principaux. Afin d'analyser l'emploi des anaphores nominales (comprenant tous les noms propres et les SN) dans les chaînes sur les personnages principaux, nous allons examiner le coefficient de stabilité (Perret, 2000). Cet indicateur se base sur le décompte des différentes désignations d'un référent par rapport au nombre total d'anaphores nominales renvoyant à ce référent. Le chiffre se présente entre 0 et 1. D'après Oberlé *et al.* (2018 : 127), 1 signifie que la chaîne est considérée comme stable du fait que les reprises nominales sont identiques dans tous les maillons alors que 0 veut dire que les reprises sont toutes différentes et que de ce fait la chaîne est instable⁴⁸.

⁴⁸ Perret (2000) calcule le coefficient de stabilité en divisant le nombre total d'anaphores nominales par le nombre

D'après nos observations du texte *Fantôme*, le nombre d'anaphores nominales de la chaîne de *fantôme* s'élève à 70 avec 26 désignations⁴⁹. La chaîne *Raoul*, elle, compte 62 anaphores nominales ayant 20 désignations différentes. Enfin, la chaîne de *Christine Daaé* comprend 49 anaphores nominales avec 22 reprises différentes. Le chiffre moyen de coefficient de stabilité des personnages principaux du roman *Fantôme* est 0,62.

Dans la chaîne de *Dong-hyuk* du roman *Arbre*, 37 désignations différentes au sein des 201 anaphores nominales se manifestent et, dans la chaîne de *Young-shin*, 34 reprises différentes se présentent parmi 138 anaphores nominales. Le coefficient de stabilité moyen du texte *Arbre* est 0,79. Le tableau suivant récapitule le coefficient de stabilité des chaînes principales de chaque roman :

	Personnage	Nbre des anaphores nominales	Nbre des désignations différentes	Coefficient de stabilité	
<i>Fantôme</i>	Fantôme	70	26	0,63	0,62
	Raoul	62	20	0,67	
	Christine Daaé	49	22	0,56	
<i>Arbre</i>	Dong-hyuk	201	37	0,82	0,79
	Young-shin	138	34	0,75	

Tableau 27 Le coefficient de stabilité des chaînes des personnages principaux

Comme nous l'avons souligné plus haut, plus le chiffre est proche de 1, plus la chaîne est stable. Les chaînes des personnages principaux du texte *Arbre* sont moyennement plus stables que celles de *Fantôme*. Cette disparité peut être, pour une part, expliquée par la fréquente redénomination *via* le nom propre du coréen.

Pour vérifier notre argument sur la fréquente redénomination du coréen, nous avons

de désignations différentes, tandis que Oberlé et *al.* (2018) choisissent une formule différente : $1-(x-1)/(n-1)$ où x est « le nombre de désignations différentes » et n « le nombre total d'anaphores nominales ». Après avoir tenté plusieurs fois les différentes façons de le calculer, nous choisissons, par commodité, la formule « $1- x/n$ » qui concorde avec les modalités de calcul de CR Viewer.

⁴⁹ Force est de constater la particularité de la chaîne de *fantôme* qui se distingue des deux autres chaînes des personnages lors de l'analyse des anaphores nominales. Les anaphores nominales de la chaîne *fantôme* ne comprennent pas les noms propres, parce que le vrai nom de ce personnage, *Erik*, n'est donné qu'à partir du chapitre 10 du roman qui ne fait pas partie de notre étude. En conséquence, la chaîne *fantôme* est exceptionnellement constituée de SN.

compté le nombre des occurrences de la redénomination *via* le nom propre⁵⁰ dans les chaînes des personnages principaux de chaque roman⁵¹. Ce mode de saisie référentielle est communément observé dans toutes les chaînes des personnages. Cependant, une nette différence proportionnelle se présente :

		<i>Fantôme</i>		<i>Arbre</i>	
		Raoul 223 occ.	Christine 128 occ.	Dong-hyuk 467 occ.	Young-shin 280 occ.
Nbre de redénomination <i>via</i> NPR		34 (15,2%)	29 (22,6%)	165 (35,3%)	110 (39,2%)
NP nu	Nom complet	-	9	7	8
	Nom de famille	-	2	3	-
	Prénom	32	10	144	96
NP modifié		2	8	11	6

Tableau 28 La redénomination *via* le nom propre

Ce tableau indique que les CR des personnages principaux du texte *Arbre* favorisent la redénomination *via* le nom propre par rapport à celles dans le texte *Fantôme* : dans le roman coréen, plus d'un tiers des occurrences de chaque personnage principal se fait par la redénomination *via* le nom propre, alors que sa proportion se situe entre 15% et 23% dans *Fantôme*.

Le coefficient de stabilité des CR des personnages principaux dans les deux romans nous permet de voir différents degrés de diversité des anaphores nominales employées et démontre en même temps que le coréen s'appuie plus sur la redénomination que le français lors de la référence aux personnages.

⁵⁰ Nous avons centré sur le nom propre parce que la reprise à l'identique par le SN n'est observée que dans la chaîne de *Raoul* dans une série de descriptions de ce personnage.

⁵¹ Le référent *fantôme* est exclu ici, parce que ce personnage n'est pas introduit par son nom comme nous l'avons mentionné dans la note précédente.

4 Synthèse

Dans ce chapitre, nous avons analysé les CR, qui sont différemment représentées dans les deux romans. Nos deux textes sélectionnés en raison de leurs similarités au niveau du volume, de l'époque de rédaction et du point de vue du narrateur manifestent cependant des disparités intéressantes liées aux caractéristiques de chaque langue et aux traits du genre textuel.

Après avoir présenté les chiffres-clés récupérés *via* le logiciel CR Viewer, nous avons voulu montrer, dans un premier temps, la façon dont on constitue une chaîne dans les deux textes : les emplois des marqueurs référentiels diffèrent selon la langue et cet aspect influence la composition des CR. Dans le texte coréen, nous avons remarqué quelques particularités qui se distinguent du texte français : l'importance du pronom zéro, l'utilisation moindre du pronom personnel et l'emploi davantage présent des marqueurs référentiels de basse accessibilité.

Les deux particularités des usages des pronoms peuvent être explicitées par quelques aspects concernant le système langagier que nous avons décrit au sein de cette thèse, notamment dans les chapitres 3 et 4. D'abord, en coréen, par une haute dépendance contextuelle, nous rencontrons souvent le phénomène d'ellipse, non seulement pour le sujet mais aussi pour l'objet. Dans de nombreux cas, les pronoms personnels ne se réalisent pas et les pronoms zéros se présentent plus souvent. Ensuite, cette langue n'a pas de système de genre, ce qui rend la reprise par le pronom personnel de la troisième personne des référents non-humains impossible. Ceux-ci sont simplement désignés par leurs noms ou par les autres formes variées de SN (déterminés par le démonstratif ou le numéral, ou bien modifiés par l'adjectif ou la relative). Ce dernier point est révélé plus clairement par notre test empirique des textes encyclopédiques sur les référents non-humains. Enfin, considérant le fait que le pronom, entre autres, de la troisième personne, est introduit par la nécessité de traduire l'anglais en coréen, au début du 20^{ème} siècle, son usage est relativement limité par rapport à celui du français.

En ce qui concerne les expressions référentielles de basse accessibilité comme le nom propre et le SN, nous avons démontré certaines conditions de leur utilisation par le biais d'exemples extraits de nos textes. Elles apparaissent dans les contextes ayant une interférence potentielle (telle que la présence des référents concurrentiels ou des tours de paroles de plusieurs personnages) où il est nécessaire de préciser le référent en question ou

dans les cas de ruptures contextuelles. Concernant ces dernières, nous avons remarqué que la langue coréenne ne réagit pas de la même façon que le français lors de certaines ruptures telles que la rupture aspectuo-temporelle et celle d'argumentation.

Selon nos observations, l'usage plus important et moins sensible des marqueurs pleins du texte coréen peut être également explicité, d'une part, par une sorte de « rétroaction » de l'utilisation moindre des pronoms que nous venons de mentionner et, d'autre part, par la façon de transmettre l'information supplémentaire sur le référent (dans de nombreux cas, la transmission de l'information est faite *via* la nominalisation, à savoir *via* la forme du nom propre modifié par une relative ou celle de descriptions définies longues).

Dans un second temps, afin d'examiner le fait que les différences inter-langues ne sont pas le seul facteur déterminant les caractéristiques des CR représentées dans notre corpus, nous avons d'abord analysé rapidement quelques textes encyclopédiques. Cette partie expérimentale indique que malgré le fait que les CR d'une langue présentent une tendance générale, elles peuvent également représenter quelques différences selon le genre du texte. Dans le cas du coréen, une de ces différences concerne la diversité lexicale des CR : le type de désignation est plus diversifié dans le roman que dans le texte encyclopédique. A partir de cette observation, nous avons ensuite décrit quelques mesures d'analyse susceptibles de montrer l'influence du genre textuel sur les CR : le nombre de singletons, la longueur et la portée des CR, le contexte spatio-temporel du roman et les divers aspects de CR des personnages tels que le type de la première mention, le niveau énonciatif de l'apparition du personnage et la diversité des anaphores nominales utilisées.

Au sein de l'interprétation de ces paramètres dans le cadre du genre textuel, nous avons remarqué que chaque roman a son propre rythme de progression et des dispositifs qui se répercutent sur les caractéristiques des chaînes. En effet, notre roman coréen a moins de mots et plus de chaînes, notamment des chaînes courtes, par rapport au roman français. Ces caractéristiques peuvent être interprétées en rapport avec la vitesse de narration : le texte coréen se caractérise par davantage de motifs variés : épisodes, sous-thèmes, circonstances, *etc.*, et le temps dans le récit passe plus vite que celui dans le roman français.

En fin de compte, nous voudrions souligner que les caractéristiques de chaque langue sont des éléments essentiels à considérer lors de l'analyse des CR. En effet, elles influencent énormément la composition des CR. Cependant, il est aussi important de regarder les traits du genre textuel concerné et la structure textuelle particulière du texte. Par cela, nous arriverons à comprendre plus subtilement, c'est-à-dire avec un point de vue multi-factoriel, les

CR et la manière dont elles contribuent à la cohérence du texte entier.

Conclusion

Tout au long de cette thèse, nous avons essayé de comprendre et d'analyser le phénomène des CR d'un point de vue comparatif entre le français et le coréen qui sont deux langues complètement différentes. En guise de conclusion, nous synthétiserons ce que nous avons découvert au sein de cette recherche.

En présentant les caractéristiques et les emplois des expressions référentielles dans la première partie de cette thèse, nous avons remarqué quelques disparités importantes entre les deux langues. La première concerne le fait que, contrairement au français, le coréen ne possède pas de système d'article. La détermination d'un nom est exprimée, de manière indirecte, par le contexte d'énonciation et par d'autres dispositifs grammaticaux tels que la particule et le prédicat. Les noms peuvent également être déterminés directement par des expressions épithètes comme le nom, la construction possessive, l'adjectif, la construction relative et la catégorie grammaticale dite *gwanhyungsa*. Ces éléments qualificatifs se placent devant le nom et lui attribuent un trait déterminatif (Hong Chai-song, 2001).

La deuxième disparité que nous souhaitons souligner est le statut différent du pronom dans les deux langues. À la différence du français, dans la langue coréenne, les pronoms, notamment les pronoms personnels de la troisième personne du singulier, ont un statut moins important. L'emploi de ces derniers ne commence à être développé qu'à partir du début du 20^{ème} siècle par la nécessité de les traduire par *he* et *she* dans les romans anglais. Tenant compte du fait que le coréen ne possède pas la conception de *genre grammatical*, les emplois de *geu* (il) et *geu nyeo* (elle) sont réservés plutôt à l'écrit et dans le cas général, la forme « *jisi gwanhyungsa* (démonstratif) + N » est plus communément utilisée pour désigner la troisième personne. Cela explique également, dans une certaine mesure, la raison pour laquelle en linguistique coréenne, ce sont les pronoms démonstratifs qui font l'objet de nombreuses recherches, alors que les pronoms personnels de troisième personne sont peu étudiés.

Nous avons également remarqué que le degré de dépendance du contexte est différent selon la langue. Le coréen semble être plus dépendant du contexte d'énonciation par rapport au français. Cela peut être expliqué par le fait que la détermination d'un nom est, pour une

part, influencée par le contexte et que les pronoms zéros s'emploient fréquemment non seulement pour le sujet mais aussi pour l'objet.

Ces différences inter-langues méritaient d'être mentionnées au préalable parce que ce sont des facteurs qui concernent certains aspects de la représentation des CR dans les deux langues. En effet, nous avons pu observer quelques phénomènes sur lesquels ces disparités se répercutent.

D'après nos observations, le pronom personnel est une des catégories grammaticales dominantes dans le texte français, alors que dans le roman coréen, c'est le pronom zéro qui est employé le plus souvent. L'utilisation des pronoms se manifeste différemment selon la langue, et, cela est étroitement liée au statut du pronom et à la dépendance contextuelle de chaque langue.

Ensuite, nous avons remarqué que l'occurrence des marqueurs pleins dans les CR montre également certaines disparités inter-langues. Premièrement, les expressions référentielles de basse accessibilité, les noms propres et les SN, sont plus souvent utilisées dans le roman coréen qu'en français. Deuxièmement, les proportions des SN et des noms propres modifiés dans les deux textes démontrent que les façons dont les informations sont transmises et dont un SN se forme sont étroitement liées l'une à l'autre et fonctionnent différemment dans les deux langues. La langue coréenne utilise plus souvent la construction relative que la langue française et cela en vue de deux objectifs : la détermination du nom et la transmission de l'information sur le référent. Enfin, notre analyse sur les contextes d'apparition des marqueurs pleins permet de voir que la contrainte telle que l'emploi de conjonction ou d'adverbe dans des phrases contiguës pour que la reprise à l'identique des marqueurs pleins soit possible n'est pas nécessairement observé dans le texte coréen.

Par ailleurs, au-delà de l'analyse basée sur les disparités inter-langues et leurs influences dans la représentation des CR, nous avons essayé de mettre en évidence le rôle du genre textuel et les traits individuels de chaque texte. L'importance du genre textuel étant traitée dans les études des CR menées par nos prédécesseurs, nous avons pu observer qu'il en est de même pour le coréen. Par exemple, dans les textes de *Wikipédia*, à la différence du roman, nous observons moins de diversité lexicale, plus de redénominations *via* marqueurs pleins dans la CR et l'absence du pronom personnel. Bien qu'un léger décalage proportionnel puisse se présenter selon le type de référent (humain / non-humain), il n'en demeure pas moins que

le genre du texte est un facteur essentiel pour expliciter la représentation des CR. La corrélation entre les caractéristiques des CR et les divers éléments constitutifs du roman tels que les cadres spatio-temporels, les incidents et les épisodes, ce que nous appelons les traits individuels de chaque texte, jouent aussi un rôle important. Ces éléments qui fonctionnent comme les moyens de découper le récit en blocs influencent la longueur et la portée des CR.

Arrivée au terme de ce parcours, nous avons conscience des limites de notre travail qui sont globalement liées à la sélection du corpus. En effet, nos analyses ont été réalisées sur la base unique du corpus constitué respectivement d'un seul texte choisi dans chaque langue. Ce choix, qui a déjà été justifié et qui semble avoir un certain mérite dans le sens qu'il rend compte des CR sur le long terme, n'est pourtant pas une option parfaite pour obtenir un résultat généralisable ou qualitatif. Au sein de l'analyse, nous avons ressenti la nécessité de compléter les données, à savoir le corpus, pour pouvoir enrichir la comparaison inter-langues. Afin de suppléer cette lacune, nous avons brièvement testé quelques exemples de texte encyclopédique, *Wikipédia*, mais, cette comparaison empirique n'est pas suffisante pour être à la hauteur du travail. Nous nous proposons donc d'envisager quelques prolongements à ce travail.

D'une part, étant donné que les romans se déroulent autour des personnages principaux et que ces derniers sont des référents les plus importants du texte, il serait intéressant d'examiner les textes dont le référent principal est une entité non-humaine pour voir si les différences au niveau des caractéristiques des CR sont observées selon le type de référent et si ces différences résultent des particularités d'un système langagier. Comme nous l'avons déjà mentionné en examinant quelques exemples de texte encyclopédique, en coréen, les pronoms personnels de la troisième personne ne s'emploient pas lors de la référence aux entités non-humaines. Diversifier le type de texte cible pourrait être un des moyens de démontrer, de manière plus efficace, des disparités entre les deux langues.

D'autre part, pour effectuer un travail rentable dans une perspective inter-langues, il serait également intéressant de constituer un corpus des textes écrits dans une autre langue et traduits en français et en coréen. Au début de ce parcours, choisir les textes traduits comme corpus n'était pas une option, parce que nous avons pensé que les tournures de la traduction manquaient de naturel et nous empêcheraient de voir les caractéristiques d'une langue en tant que telles. C'est la raison pour laquelle nous avons choisi respectivement un roman de chaque langue. Cependant, une des difficultés que nous avons rencontrée au cours de

l'analyse concerne les différences du thème, du type ou du nombre des motifs, de la vitesse de narration ou de la proportion du discours rapporté de chaque texte qui s'avèrent plus grandes que nous avons pu imaginer même si les deux romans partagent quelques points en commun tels que le point de vue omniscient et l'époque de la rédaction. De ce fait, lors de l'analyse, nous avons essayé de prendre en compte ces traits individuels de chaque texte mais il était parfois difficile de trouver les contextes similaires pour expliciter tous les cas traités. Comparer le même texte traduit respectivement dans chaque langue pourrait résoudre ce problème et nous permettrait de voir les caractéristiques des CR dans les mêmes conditions contextuelles.

Bien que cette thèse laisse certaines limites, elle représente certainement une originalité qui consiste à avoir fourni une perspective théorique sur les expressions référentielles, et, aussi à avoir essayé de comparer les CR en français et en coréen. Pour autant que nous le sachions, notre thèse est la première étude comparative du domaine des CR entre ces deux langues. En linguistique coréenne notamment, il s'agit d'une tentative inédite non seulement parce que le sujet est original mais aussi parce qu'elle apporte de nouveaux outils méthodologiques. Nous espérons donc que cette étude pourrait fournir une donnée de base pour ceux qui souhaitent mener une recherche sur les CR en coréen et pour ceux qui travaillent dans le domaine de la référence d'un point de vue comparatif entre le français et le coréen. En outre, considérant que les CR est un des moyens importants dans la production et la compréhension du texte, notre thèse pourrait également servir à l'enseignement du français pour les apprenants coréens et celui du coréen pour les Français qui apprennent le coréen.

En tenant compte de nos limites et de nos avantages, nous espérons continuer ultérieurement cette série de recherches contrastives à portée étendue en diversifiant le type de texte étudié et la stratégie d'analyse.

Bibliographie

- ACHARD-BAYLE G. (1998). « Référence, identité, changement : la désignation des référents en contextes évolutifs. Etudes de cas : les récits de métamorphoses », *L'Information Grammaticale*, 77, p.50-53.
- ADAM J. M. (2005), *La linguistique textuelle : introduction à l'analyse textuelle des discours*. Paris : Armand Colin.
- AHN J. K. & EUN H. Y. (2006). « 영어와 한국어 지시대명사 대조 연구 (A Contrastive Study of Korean and English Demonstratives) », *The new studies of English language & literature*, 35, p.229-249.
- APOTHELOZ D. (1995). *Rôle et fonctionnement de l'anaphore dans la dynamique textuelle*, Genève : Droz.
- ARIEL M. (1990). *Accessing noun-phrase antecedents*, Londres : Routledge.
- BAE D. H. (1998). « 한국인 일본어 학습자를 대상으로한 일본어 지시표현에 지도에 관한 연구 (A Teaching Method of Japanese Indication for Korean Student) » *Sangmyung Inmungwahakyeongu*, 7, p.1-20.
- BARBET C. & DE SAUSSURE L. (2012). « Présentation: Modalité et évidentialité en français », *Langue française*, 173, p.3-12.
- BAUMER E. (2015). *Noms propres et anaphores nominales en anglais et en français : étude comparée des chaînes de référence*, Paris : L'Harmattan.
- BAUMER E. (2017). « Chaînes de référence et point de vue dans la fiction littéraire : le cas des nouvelles courtes », *Langue française*, 195, p.73-90.
- BAUMER E., DIAS D. & SCHNEDECKER C. (2020). « Les chaînes de référence dans un corpus contrastif (allemand–anglais–français) de romans réalistes du XIXe siècle : analyse quantitative et qualitative », *Phraséologie et stylistique de la langue littéraire*, Berlin : Peter Lang, p.211-229.
- BAUMER E., DIAS D., GARDELLE L. & PRAK-DERRINGTON E. (2021). « Quelles régularités pour les chaînes de référence dans le genre 'nouvelles d'actualité' ? Exploration contrastive allemand / anglais / français », *Travaux de linguistique*, 82, p.17-43.
- BENVENISTE E. (1966). *Problèmes de linguistique Générale*, Paris : Gallimard.
- CAPIN D. (2014). « Chaînes de référence dans les textes médiévaux non-narratifs : les Year Books ou

l'élaboration d'une écriture juridique », *Langages*, 195, p.61-78.

CAPIN D., GLIKMAN J., SCHNEDECKER C. & TODIRASCU A. (2021). « Le rôle des chaînes de référence dans la structuration textuelle : étude diachronique de l'ancien français au français moderne », *Langages*, 224(4), p.87-107.

CHAROLLES M. (1978). « Introduction aux problèmes de la cohérence des textes : (Approche théorique et étude des pratiques pédagogiques). », *Langue française*, (38), p.7-41.

CHAROLLES M. (1987). « Contraintes pesant sur la configuration des chaînes de référence comportant un nom propre ». *Cahiers du centre de recherches sémiologiques de Neuchâtel*, 53, p.29-55.

CHAROLLES M. (1988). « Les plans d'organisation textuelle : périodes, chaînes, portées et séquences », *Pratiques*, 57(9), p.3-13.

CHAROLLES M. (1995a). « Cohésion, cohérence et pertinence du discours », *Travaux de Linguistique*, 29, p.125-151.

CHAROLLES M. (1995b). « Comment repêcher les derniers ? Analyse des expressions anaphoriques en 'ce dernier' », *Pratiques : linguistique, littérature, didactique*, 85, p.89-113.

CHAROLLES M. (1997). « L'encadrement du discours : Univers, champs, domaines et espaces », *Cahier de Recherche Linguistique*, 6, p.1-73.

CHAROLLES M. (2001). « De la phrase au discours : quelles relations », *La sémantique des relations* (in A. Rousseau éd.), Lille : Université de Lille III, p.237-260.

CHAROLLES M. (2002). *La référence et les expressions référentielles en français*, Paris : Ophry.

CHAROLLES M. (2006). « Cohérence, pertinence et intégration conceptuelle », *Des discours aux textes : modèles et analyses* (in Lane P. éd.), Rouen : Presses Universitaires de Rouen, p.39-74.

CHAROLLES M. (2011). « Cohérence et cohésion du discours », *Dimensionen der Analyse Texten und Diskursivent – Dimensionen dell'analisi di testi e discorsi* (In K.Hölker & C. Marelllo ed), p.153-173.

CHANG K. H. (2002). « 대명사 (Les Pronoms) », *Saegugeosanghwal*, 12-2, p.147-161.

CHANG K. H. (2004). « 국어 지시 표현의 유형과 성능 (Types and functions of Korean referring expressions) », *Korean semantics*, 15, p.51-70.

CHASTAIN C. (1975). « Reference and context », *Language, Mind and Knowledge* (in K. Gunderson ed.), Minneapolis : University of Minnesota Press, p.194-269.

- CHOI J. J. (2016). « 명사구와 관련된 정보구조와 한정성의 상관관계 (La corrélation entre la détermination et la structure informationnelle du SN) », *Korean semantics*, 54, p. 151-173.
- CHOI-JONIN I. (2008). « Particles and postpositions in Korean », *Adpositions. Pragmatic, semantic and syntactic perspectives* (in D. Kurzon & S. Adler éd.), Philadelphia : J. Benjamins Publishing Company, p. 133-170.
- COL G. (2012). « Focalisation, saillance et instruction de construction du sens », *Focalisation(s) : saillance dans les langues : lexicale, syntaxe, prosodie* (In P. Cappeau & S. Hanote éd.), Rennes : Presses Universitaires de Rennes, p.84-101.
- COMBETTES B. & TOMASSONE R. (1988). *Le texte informatif : aspects linguistiques*, Bruxelles : De Boeck.
- COMBETTES B. (1988a). *Les constructions détachées en français*. Paris : OPHRYS.
- COMBETTES B. (1988b). « Fonctionnement des nominalisations et des appositions dans le texte explicatif », *Pratiques*, 58.1: 107-119.
- COMBETTES B. (2006). « Cataphore et texte littéraire : aspects diachroniques ». *Langue littéraire et changements linguistiques* (in F. Berlan éd.), Paris : Presses de l'Université Paris-Sorbonne, p.385-396.
- CORBLIN F. (1985a). « Remarques sur la notion d'anaphore », *Revue québécoise de linguistique*, 15 (1), p.173–195.
- CORBLIN F. (1985b). « Les chaînes de référence : analyse linguistique et traitement automatique », *Intellectica*, 1(1), p.123-143.
- CORBLIN F. (1995). *Les formes de reprise dans le discours. Anaphores et chaînes de référence*, Rennes : P. U. Rennes.
- CORNISH F. (2000). « L'accessibilité cognitive des référents, le centrage d'attention, et la structuration du discours : une vue d'ensemble », *Verbum*, 22(1), p.7-30.
- CORNISH F. (2011). « Prédication, focalité, topicalité et énoncés thétiques : une description GFD. », *Fonctionnalisme et description linguistique* (in M. Jadir éd.), p.131-156.
- DENIS D. & SANCIER-CHATEAU A. (1994). *Grammaire du français (Le livre de poche)*, Paris : Librairie générale française.
- DUPONT V. & BESTGEN Y. (2006). « Learning from technical documents: The role of intermodal referring expressions », *Human factors*, 48(2), p.257-264.
- DUVAL M. (2003). *Le problème de l'interrogation indirecte totale d'après ses marques (en français)*,

anglais et coréen) : étude contrastive et typologique, Thèse de doctorat, Université Paris-Sorbonne.

FÈVRE-PERNET C. & ROCHÉ M. (2005). « Quel traitement lexicographique de l'onomastique commerciale ? Pour une distinction Nom de marque/Nom de produit », *Corela. Cognition, représentation, langage*, (HS-2).

FRANCKEL J.-J. (1998). « Référence, référenciation et valeurs référentielles », *Sémiotiques*, 15, p.61-84.

FREGE G. (1892). *Ecrits logiques et philosophiques*, Paris : Seuil.

FREYERMUTH S. (2000). « L'économie de la reprise anaphorique : un révélateur de compétences stylistiques dans les écrits d'élèves en échec scolaire », *Répétition, altération, reformulation*, p.173-187.

GALMICHE M. (1983). « Les ambiguïtés référentielles ou les pièges de la référence », *Langue française*, 57, p.60-86.

GARDELLE L., ROSSI C. & VINCENT-DURROUX L. (2019). « La gestion de l'anaphore en discours: complexités et enjeux », *Cahiers de praxématique*, 72, p.1-17.

GARY-PRIEUR M.-N. (1991). « Le nom propre constitue-t-il une catégorie linguistique ? », *Langue française*, 92, p.4-25.

GARY-PRIEUR M.-N. & NOAILLY M. (2019). « Le nom propre », in *Encyclopédie grammaticale du français*, <www.encyclogram.fr>.

GENOT G. & LARIVAILLE P. (1984). « Style narratif, rhétorique, tradition (sur le Novellino) », *Revue Romane*, 19, p.284-297.

GIVON T. (1983). Topic continuity in discourse : An introduction, *Topic continuity in discourse. A quantitative cross language study* (in T. Givon ed.), Amsterdam: J. Benjamins Publishing Company.

GJESDAL A. (2008). *Etude sémantique du pronom ON dans une perspective textuelle et contextuelle*, Thèse de doctorat, Université de Bergen.

GLIKMAN J., GUILLOT-BARBANCE C. & OBRY V. (2014). « Les chaînes de référence dans un corpus de textes narratifs médiévaux : traits généraux et facteurs de variation », *Langages*, 195, p.43-60.

GROSS G. (1996). *Les expressions figées en français : noms composés et autres locutions*. Paris : Ophrys.

- GUÉRIN O. (2014). « Construction du référent, textualité et genre discursif : les anaphores génériques dans les séquences encyclopédiques des récits de voyage », *In : SHS Web of Conferences. EDP Sciences*, p.3091-3110.
- GUILLOT C. (2006). « Démonstratif et déixis discursive : analyse comparée d'un corpus écrit de français médiéval et d'un corpus oral de français contemporain », *Langue française*, 152, p.56-69.
- GUO C. (2021). « Chaînes de référence dans les textes encyclopédiques : étude contrastive en français et en chinois », *Travaux de linguistique*, (1), p.129-152.
- HALLIDAY M.A.K & HASAN R. (1976). *Cohesion in English*, Londres : Longman.
- HEO J. Y. (2017). « 국어사 시대 구분과 20세기 초의 문법 변화 (The division of Korean language history and the grammatical change in the early 20th century) », *Gugjeonmun*, 73, p.7-32.
- HONG C. S. (2001). « 한국어의 명사 I (Les noms coréens I) », *Gugeohakui ihae*, p.129-144.
- HONG J. S. (2016). « 한국어사에서 20세기 초 한국어의 위상과 문법 특징 (Historic status and grammatical characteristics of Korean language in the early 20th century) », *Korean linguistics*, 71m p.1-22.
- HONG Y. T. (2009). « 프랑스어와 한국어의 명사구 (Syntagmes nominaux du français et du coréen) ». *Enseignement de langue et littérature françaises*, 31, p.243-280.
- HOU J. (2021). *Étude de la saillance référentielle dans un corpus de textes narratifs : analyse contrastive des chaînes de référence en français et en chinois*, Thèse de doctorat, Université Sorbonne Nouvelle – Paris 3.
- HU J. (2011). « 한국어와 중국어 지시어에 관한 대조 연구 (Contrastive analysis of korean and chinese Deictic terms) », Thèse de doctorat, Hanyang University.
- HUSSON-CASTA I. (1997). *Le travail de «l'obscur clarté» dans "Le fantôme de l'Opéra" de Gaston Leroux*. Paris : Lettres modernes,
- JONASSON K. (1994). *Le nom propre*, Louvain-la-Neuve : Duculot.
- KIM K. S. (2010). « 국어 3인칭 대명사 ‘그’ : 프랑스어, 영어와 비교연구 (Le pronom *keu* en coréen : en comparaison du français et de l'anglais) », *Enseignement de langue et littérature françaises*, 34, p.101-121.
- KIM K. S. (2017). « 지시 연쇄 삼각형: 조음과 연상, 직관적 동일 지시 (The Triangle of Referential Chains: Anaphor, Association and Intuitive Equi-reference) », *Language Research*, 53.3, p.391-414.

- KIM M. H. (2011). « 조사 이,가 와 은,는 의 기본전제와 기능 분석 (An analysis of the basic premises and functions of Korean particles *i/ka* et *eun-neun*) », *Discourse and Cognition*, 18, p.23-64.
- KIM S. Y. (2010). *Propriétés syntaxico-sémantiques des verbes à complément en -e en coréen*, Thèse de doctorat, Université Paris-Est.
- KLEIBER G. (1981). *Problèmes de référence : descriptions définies et noms propres*, Paris : Klincksieck.
- KLEIBER G. (1983). « Métaphores et vérité », *Linx*, 9.1, p.89-130.
- KLEIBER G. (1994). *Anaphores et pronoms*, Louvain-la-Neuve : Duculot.
- KLEIBER G. (1997). « Sens, référence et existence : que faire de l'extra-linguistique ? », *Langages*, 127, p.9-37.
- KLEIBER G. (2001). *L'anaphore associative*, Paris : PUF.
- KLEIBER G. (2003). « Adjectifs démonstratifs et point de vue », *Cahiers de praxématique*, 41, p.33-54.
- KLEIBER G. (2017). « De la complexité de la personne ou Sur l'irréductibilité de JE et Tu », *PIE Peter Lang*, p.151-167.
- KURODA S.-Y., BEFFA M.-L. & BOREL M. (1973). « Le jugement catégorique et le jugement thétique : exemples tirés de la syntaxe japonaise », *Langages*, 30, p.81-110.
- KUYUMCUYAN A. (2018). « L'extension des pseudo-clivées », *SCOLIA Revue de linguistique*, 32, p.95-116.
- KWON S.-N. & ZRIBI-HERTZ A. (2006). « Pluriel et généricité nominale : les enseignements du coréen. Noms nus et généricité », *Noms nus et généricité*, Saint-Denis : Presses Universitaires de Vincennes, p.159-195.
- LANDRAGIN F. & TANGUY N. (2014). « Référence et coréférence du pronom indéfini on », *Langages*, 195, p.99-115.
- LANDRAGIN F. (2011). « De la saillance visuelle à la saillance linguistique », *Aspects linguistiques et communicatifs de la mise en évidence dans un texte*, Presses Universitaires de Franche-Comté, p.67-84.
- LANDRAGIN F. (2012). « La saillance : questions méthodologiques autour d'une notion multifactorielle », *Faits de langues*, 39(1), p.15-31.
- LANDRAGIN F. (2017). « Analyse, visualisation et identification automatique des chaînes de coréférences : des questions interdépendantes ? », *Langue française*, 195, p.17-34.

- LARSSON B. (1997). *Le bon sens commun*, Lund : Lund University Press.
- LEE J. T. (2011). « 고유 명사와 보통 명사 - 분류 기준과 개념 설정을 중심으로 (Proper noun and common noun – centered on classification and their concepts) », *Grammar education*, 15, p.267-283.
- LEROY S. (2004). *Le nom propre en français*, Paris : OPHRYS.
- LONGO L. (2013). *Vers des moteurs de recherche « intelligents » : un outil de détection automatique de thèmes. Méthode basée sur l'identification automatique des chaînes de référence*, Thèse de doctorat, Université de Strasbourg.
- LONGO L. & TODIRASCU A. (2010). « Genre-based Reference Chains Identification for French », *Investigationes Linguisticae*, Volume XXI, pp. 57-75.
- LONGO L. & TODIRASCU A. (2014). « Etude des chaînes de référence dans des textes non-narratifs », *Langages*, 195, 79-98.
- LUNDQUIST L. (2005). « Noms, verbes et anaphores (in)fidèles. Pourquoi les Danois sont plus fidèles que les Français ? », *Langue française*, 145, p.73-91.
- LUNDQUIST L., MINEL J-L. & COUTO J. (2012). « La navigation discursive. L'anaphore résomptive et mouvement discursif », *L'analyse du discours dans la société*, Honoré Champion, p.365-389
- MAES A., ARTS A. & NOORDMAN L. (2004). « Reference Management in Instructive Discourse », *Discourse Processes*, 37/2, p.117-144.
- MAILLARD M. (1974). « Essai de typologie des substituts diaphoriques », *Langue française*, 21, p.55-71.
- MAINGUENEAU D. (1994). *Syntaxe du français*, Paris : Hachette Supérieur.
- MANUÉLIAN H. & SCHNEDECKER C. (2021). « Introduction des référents et premier maillon de suites coréférentielles : la réalité des textes », *Langages*, 224(4), p.67-86.
- MILNER J-C. (1976). « Réflexions sur la référence », *Langue française*, 30, p.63-73.
- MILNER J.C. (1982). *Ordres et raisons de langue*, Paris : Seuil.
- MOESCHLER J. & REBOUL A. (1994). *Dictionnaire encyclopédique de pragmatique*, Editions du Seuil.
- MOLINO J. (1982). « Le nom propre dans la langue », *Langages*, 66, p.5-20.
- NAM K. S., KO Y. K., YOO H. K., & CHOI H. Y. (2019). *새로 쓴 표준 국어문법론 (Nouvelle Grammaire standard du coréen)*, Seoul : Hangugmunhwasa.

- NEVEU F. (2000). « L'apposition : concepts, niveaux, domaines », *Nouvelles recherches sur l'apposition*, p.3-17.
- OBERLE B., SCHNEDECKER C., BAUMER E., CAPIN D., GLIKMAN J., GUO C., REVOL T., TODIRASCU A. & TUSHKOVA J. (2018). « Les chaînes de référence dans les textes encyclopédiques du 12e au 21e siècle : étude longitudinale », *Travaux de linguistique*, 77, p.67-141.
- OBERLE B. (2018). « SACR: A drag-and-drop based tool for coreference annotation », In: *Conference: Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*.
- OBRY V., GLIKMAN J., GUILLOT-BARBANCE C. & PINCEMIN B. (2017). « Les chaînes de référence dans les récits brefs en français : étude diachronique (XIIIe-XVIe s.) », *Langue française*, 195, p.91-110.
- OH C. Y. (2012). « 국어의 명사 한정 체계(Le système de détermination nominale en coréen) », *Eomunhak*, 117, p.69-99.
- PARK D. Y. (2001). « 불어와 한국어의 명사 한정-관사 UN, LE 와 조사 {이/가},{은/는} 의 의미연구 (Détermination nominale en français et en coréen) », *Enseignement de langue et littérature françaises*, 11, p.121-142.
- PARK H. S. (2014). « 프랑스어 관사와 한국어 조사 {-가, -는, -를} 에 관한 비교 연구: 표현적 가치를 중심으로 (Une étude comparative entre les articles français et les particules coréenne '-ga, -nun, -lul' – autour de la valeur expressive) », *Études de la culture française et des arts en France*, 48, p.159-198.
- PARK H. S. (2017). « 프랑스어 3인칭 대명사의 의미와 번역 (Le sens et la traduction des pronoms de 3ème personne du français) », *Enseignement de langue et littérature françaises*, 57, p.99-136.
- PARK Y.-O. (2015). *L'ordre des éléments de la phase en coréen : esquisse de syntaxe énonciative*, Thèse de doctorat, Université de Toulon.
- PATRON S. (2005). « Le narrateur et l'interprétation des termes déictiques dans le récit de fiction », *Studia Romanica Tartuensia*, IVa, p.187-202.
- PERRET M. (2000). « Quelques remarques sur l'anaphore nominale aux 14^e et 15^e siècles », *L'information grammaticale*, 87, p.17-23.
- REBOUL A. (1991). « Le système des pronoms personnels en français contemporain : le problème de la référence », *Revue Tranel (Travaux neuchâtelois de linguistique)*, 17, p.123-146.

- REBOUL A. (1997). « (in) cohérences et anaphores : Mythes et réalités », *Relations anaphoriques et (in)cohérences, Rodopi*, p.297-314.
- RECASENS M., HOVY E. H. & MARTÍ M. A. (2010). « A Typology of Near-Identity Relations for Coreference (NIDENT) ». In: *Conference: Proceedings of the International Conference on Language Resources and Evaluation (LREC 2010)*.
- REINHART T. (1980). « Conditions for text coherence », *Poetics today*, 1(4), p.161-180.
- RIEGEL M., PELLAT J-C. & RIOUL R. (2018). *Grammaire méthodique du français (7e édition)*, Paris : PUF.
- RIOUL R. (1983). « Les appositions dans la grammaire française », *L'information grammaticale*, 18.1, p.21-29.
- ROUQUIER M. (2018). « Les constructions clivées », in *Encyclopédie grammaticale du français*, <www.encyclogram.fr>.
- RUWET N. (1983). « Du bon usage des expressions idiomatiques dans l'argumentation en syntaxe générative », *Revue québécoise de linguistique*, 13, p.9-145.
- SALLES M. (2006). « Cohésion-cohérence : accords et désaccords », *Corela, Cognition, représentation, langage*, (HS-5).
- SCHNEDECKER C. & LANDRAGIN F. (2014). « Les chaînes de référence : présentation », *Langages*, 195, p.3-22.
- SCHNEDECKER C. & LONGO L. (2012). « Impact des genres sur la composition des chaînes de référence : le cas des faits divers », *Actes du Congrès Mondial de Linguistique Française*, p.1957-1972.
- SCHNEDECKER C., GLIKMAN J. & LANDRAGIN F. (2017). « Les chaînes de référence : annotation, application et questions théoriques », *Langue française*, 195, p.5-16.
- SCHNEDECKER C. (1997). *Nom propre et chaînes de référence*, Paris : Klincksieck.
- SCHNEDECKER C. (2005). « Les chaînes de référence dans les portraits journalistiques : éléments de description », *Travaux de linguistiques*, 51, p.85-133.
- SCHNEDECKER C. (2006a). *De l'un à l'autre et réciproquement... Aspects sémantiques, discursifs et cognitifs des pronoms anaphoriques corrélés l'un/l'autre et le premier/le second*, Bruxelles : Duculot.
- SCHNEDECKER C. (2006b). « Anaphores prédicatives démonstratives : de la cohésion syntagmatique à la cohérence textuelle », *Corela, Cognition, représentation, langage*, (HS-5).

- SCHNEDECKER C. (2011). « La notion de saillance : problèmes définitoires et avatars », *Saillance : Aspects linguistiques et communicatifs de la mise en évidence dans un texte* (O.Inkova éd.), Besançons : PUFC, p. 23-43.
- SCHNEDECKER C. (2014). « Chaînes de référence : variations selon le genre », *Langages*, 195, p.23-42.
- SCHNEDECKER C. (2017). « Les chaînes de référence : une configuration d'indices pour distinguer et identifier les genres textuels », *Langue française*, 195, p.53-72.
- SCHNEDECKER C. (2019). « De l'intérêt de la notion de chaîne de référence par rapport à celles d'anaphore et de coréférence », *Cahiers de praxématique*, 72.
- SCHNEDECKER C. (2021). *Les chaînes de référence en français*, Paris : Ophrys.
- SEARLE J. (1972), *Les actes de langage*, Paris : Hermann.
- SEO N. R. (2014). « 한국어 지시어 "이", "그", "저"와 일본어 지시어 "こ", "そ", "あ"의 비교, 대조: 화시적 지시에 한하여 (Comparative study on Korean referring expressions '이, 그, 저' and those of Japanese 'こ, そ, あ' – deictic usages) » *Gwanakeomunyeongu*, 39, p.251-268.
- SHIM B. S. (1996). « 불어와 한국어에서 '지시한정사+ 명사' 용법의 비교연구 (Étude comparative de l'emploi de 'démonstratif + N' en français et en coréen) ». *Enseignement de langue et littérature françaises*, 4, p.1-29.
- SON H. (2005), « Adverbial en MANEY en coréen : aspect, contexte et présupposition », *Cahiers de linguistique – Asie orientale*, vol. 34 2, p.239-271.
- SONG J. Y. (2011). « 영어와 한국어 지시어의 용법 차이 연구 (A contrastive study on demonstratives in English and Korean) », *Journal of language sciences*, 18-1, p.91-108.
- TUTIN A. (2002). « A corpus-based study of pronominal anaphoric expressions in French », *Proceedings of DAARC 2002 (Discourse Anaphora and Anaphora Resolution)*, Lisbon, p.18-20.
- VELUPILLAI V. (2012). *An introduction to linguistic typology*, Philadelphia : J. Benjamins Publishing Company.
- WALKER M., JOSHI A. & PRINCE E. (1998). « Centering in Naturally Occurring Discourse : An Overview » *Centering Theory in Discourse* (in WALKER M. et al. éd.), Clarendon Press, p.1-28.
- WEBBER B. L. (1988). « Tense as discourse anaphor », *Computational Linguistics*, 14(2), p.61-73.

- YANG Y. J. (2018). « 영어와 한국어의 지시표현 사용에 대한 연구 (A study on using referring expressions in English and Korean) », *The journal of linguistics sciences*, 85, p. 191-209.
- YOO H. K., SEO S. K., HAN Y. K., KANG H. H., KO S. J. & CHO T. R. (2015), *우리말 연구의 첫걸음 (Premier pas dans l'étude de notre langue)*, Seoul : Bogosa.
- YUN H. S. (2004), « Comment analyser GN-EUL *gajigo* en coréen ? », *Eoneoyeongu*, 20, p.65-73.
- ZHU W. (2018). « 말뭉치 기반 한국어 지시관형사와 중국어 지시대사 대조 연구- '이, 그, 저'와 '저 (這), 나 (那)'를 중심으로 (Contrastive Study between Demonstrative Prenouns and Demonstrative Pronouns Based on corpus - For '이, 그, 저' and '這, 那'-) », *The journal of Humanities Studies*, 110, p.25-46.

Les expressions référentielles et les chaînes de référence en français et en coréen

Résumé

Cette thèse porte essentiellement sur une description des expressions référentielles et de la chaîne de référence, et en propose une analyse contrastive en français et en coréen. La notion de chaîne de référence, qui se situe dans le domaine de la référence et de la linguistique textuelle, se présente différemment selon le genre textuel, l'époque et la langue de la rédaction. Notre objectif est de montrer la manière dont les chaînes de référence se construisent et fonctionnent dans un corpus inter-langues français-coréen, notamment au sein des deux romans, *Le fantôme de l'Opéra* (1910) et *L'arbre à feuillage persistant* (1935). À travers cette étude, nous mettons en relief non seulement les disparités inter-langues et leurs influences dans la représentation des chaînes de référence, mais aussi le rôle du genre textuel et des traits individuels de chaque texte.

Mots clés : Chaîne de référence, référence, cohérence, cohésion, anaphore, coréférence, expressions référentielles, approche inter-langue.

Résumé en anglais

This dissertation essentially aims to describe referential expressions and the reference chain and offers a contrastive analysis in French and Korean. The notion of reference chain which lies in the field of reference and textual linguistics presents itself differently depending on the textual genre, the period and the written language. Our objective is to show how reference chains are constructed and function in a French-Korean inter-language corpus, particularly in two novels, *Le fantôme de l'Opéra* (1910) et *L'arbre à feuillage persistant* (1935). Through this study, we highlight not only the inter-language disparities and their influences in the representation of reference chains but also the role of textual genre and the individual features of each text.

Keywords : Reference chain, reference, consistence, cohesion, anaphor, coreference, referential expressions, inter-language approach