



**HAL**  
open science

## Sur la canonicité en morphologie: perspective empirique, formelle et computationnelle

Géraldine Walther

► **To cite this version:**

Géraldine Walther. Sur la canonicité en morphologie: perspective empirique, formelle et computationnelle. Linguistique. Université Paris 7 Denis Diderot, 2013. Français. NNT: . tel-01535976

**HAL Id: tel-01535976**

**<https://theses.hal.science/tel-01535976>**

Submitted on 9 Jun 2017

**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.

Copyright

UNIVERSITÉ PARIS DIDEROT (PARIS 7)  
École doctorale de Sciences du Langage 132  
U.F.R. de Linguistique

Numéro attribué par la bibliothèque

--	--	--	--	--	--	--	--	--	--

## THÈSE

Nouveau régime

Pour obtenir le grade de

Docteur en Sciences du Langage

Discipline : Linguistique théorique, descriptive et automatique

Présentée et soutenue publiquement

par

**GÉRALDINE WALTHER**

Le 11 juin de l'an MMXIII

# SUR LA CANONICITÉ EN MORPHOLOGIE

Perspective empirique, formelle et computationnelle

Directeurs de thèse :

**ANNE ABEILLÉ & OLIVIER BONAMI**

Composition du jury :

Anne ABEILLÉ	LLF, Université Paris Diderot (directrice)
Olivier BONAMI	LLF, Université de Paris–Sorbonne (directeur)
Greville G. CORBETT	Surrey Morphology Group, University of Surrey (pré-rapporteur)
Bernard FRADIN	LLF, CNRS (président)
Nabil HATHOUT	CLLE-ERSS, CNRS (pré-rapporteur)
Gregory T. STUMP	University of Kentucky



À ma grand-mère, équilibre discret de la famille, beaucoup trop tôt disparue  
À mon grand-père, mon premier modèle  
À mes parents



# Remerciements

Je tiens à exprimer ici ma gratitude envers les nombreuses personnes sans lesquelles cette thèse n'aurait pas été ce qu'elle est devenue aujourd'hui.

Tout d'abord, ma directrice Anne Abeillé qui a accepté, du jour au lendemain, de me diriger, alors même qu'elle ne me connaissait pas. À travers les années qu'a duré l'élaboration de cette thèse, elle est restée un soutien indispensable. En particulier, c'est elle qui a su me jeter dans les bras d'Olivier lorsque mon sujet de recherche s'est finalement réorienté vers la morphologie. Même après mon changement de sujet, elle a toujours été là, discrète, mais présente tant dans les questions immédiatement liée à la vie de thésarde que dans celles de la vie en général.

J'ai rencontré mon co-directeur Olivier Bonami en 2010, lors d'une réunion du projet ANR PerGram. J'avais initié une réorientation complète de mon sujet et Anne lui avait confié la responsabilité de veiller sur ce nouveau projet. J'ai trouvé en lui un directeur exceptionnel ! Il m'a dirigée en sachant me donner à la fois la liberté, les conseils et les directives dont j'avais besoin pour mener à bien ce projet. Je le remercie particulièrement pour sa très grande disponibilité, sa patience aussi dans les moments où un chapitre se faisait attendre ou que je lui rapportais un énième projet ou sujet dans lequel j'étais en train de me lancer en complément du sujet principal de ma thèse. Je remercie aussi Olivier pour son soutien infailible, sa grande sensibilité et compréhension et enfin les très nombreuses discussions passionnantes qui ont ponctué ces trois dernières années : j'espère que nous en aurons encore beaucoup dans l'avenir !

Je remercie les membres de mon jury qui ont courageusement accepté de braver la lecture de ces bien nombreuses pages. Les remarques de mes deux rapporteurs, Greville G. Corbett et Nabil Hathout, m'ont été très utiles pour améliorer — du moins je l'espère — certains passages de ce travail. Merci aussi à Bernard Fradin pour ses commentaires et remarques pertinentes et encourageantes au fur et à mesure de sa lecture. Les échanges que j'ai pu

---

avoir avec Gregory T. Stump ont également été à la fois extrêmement enrichissants et encourageants. Je l'en remercie très chaleureusement.

Je suis également redevable à Bernard Colombat et Jacques Jayez qui, à l'ENS, ont veillé sur mes premiers pas dans le domaine de la linguistique et m'ont aidée à trouver ma voie entre Lyon, Heidelberg et Paris. Bernard Colombat a toujours su me rassurer, m'encourager et me guider dans mes choix d'études. Sans Jacques Jayez, qui m'a grandement soutenu et aidé dans la constitution de mon dossier d'AMN et qui a su me mettre en rapport avec Anne Abeillé, je ne serais pas aujourd'hui en train d'écrire les remerciements de cette thèse. Merci de tout cœur à eux deux !

Je ne remercierai jamais assez Jean Lowenstamm, alors directeur de l'École doctorale, pour ses interventions au moment de mon inscription en thèse à l'Université Paris Diderot. C'est lui plus que quiconque qui par son calme, sa compréhension et ses encouragements ainsi que par sa persévérance m'a aidé à passer ce moment difficile. Sans lui, je ne serai pas ici aujourd'hui. Merci.

Merci aussi à Benoît Sagot et Pollet Samvelian qui m'ont permis de m'ouvrir à des thématiques nouvelles qui ont abouti en la redéfinition de mon sujet de thèse. C'est en effet en travaillant sur la morphologie du persan pour le développement du lexique électronique PerLex avec Benoît pour le projet ANR PerGram dirigé par Pollet que s'est confirmé mon intérêt pour la morphologie.

Je remercie également Pollet Samvelian de m'avoir confrontée aux questions épineuses et passionnantes de la morphologie kurde qui auront su trouver une place dans cette thèse. Je lui dois bien d'autres remerciements encore, et pas le moindre pour m'avoir recrutée pendant six mois au sein de son projet ANR PerGram.

Cette thèse n'aurait par ailleurs jamais pris la forme qu'elle a aujourd'hui sans les collaborations que j'ai pu avoir ces dernières années. Je tiens donc à remercier de tout cœur mes co-auteurs Pegah Faghiri, Karën Fort, Lionel Nicolas, Pollet Samvelian et surtout Maris Camilleri et Benoît Sagot. Travailler avec vous fait partie des meilleurs moments que j'ai pu vivre pendant cette thèse !

Plus encore, je remercie Benoît Sagot qui a suivi, à travers nos collaborations, l'avancée de mes travaux. Notre collaboration a joué un rôle clef dans l'évolution du sujet de cette thèse. C'est grâce à elle que j'ai eu l'occasion de découvrir le domaine du Traitement Automatique des Langues. C'est notre travail commun sur l'adaptation de son formalisme lexical Alexina au modèle formel de la flexion élaboré dans cette thèse qui a permis de faire de  $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$  un modèle implémenté, de développer les mesures d'évaluation descriptive et par conséquent d'effectuer les différentes expériences décrites dans les deux derniers chapitres. Mais au

---

delà de notre collaboration, c'est à partir de nos discussions et en réponse à nombreux de ses riches et pertinents commentaires que s'est construite ma conception du lexique et de son rôle par rapport au reste de la grammaire. C'est en tant que collègue et collaborateur principal que je le remercie ici de m'avoir fait découvrir tant de choses et de m'avoir ouvert tant d'horizons, mais également de m'avoir, par son intérêt pour mon travail, encouragé plus que quiconque à le mener à bien.

Je suis également grandement redevable à Gilles Boyé, Noam Faust, et Nicolà Lampitelli pour leurs commentaires au moment de la rédaction de la partie sur l'interface morphologie-phonologie au chapitre 6 de cette thèse. Leurs conseils m'ont été très précieux et ont, j'en suis sûre, grandement amélioré la qualité de cette section. Les défauts qui lui restent sont évidemment de mon entière responsabilité.

Merci à Benoît et Maman pour avoir fièrement bataillé les nombreuses typos et en avoir éradiqué un nombre conséquent. Celles qui restent sont le résultat de leur effectif pour trop important et j'en suis naturellement seule responsable !

Le contenu de cette thèse a par ailleurs grandement bénéficié du séjour que j'ai eu la chance de faire au *Surrey Morphology Group* de l'Université de Surrey en janvier et février 2012. J'aimerais ici très profondément remercier Grev Corbett et Dunstan Brown pour leur accueil plus que chaleureux, et leurs conseils inestimables. Merci à toute l'équipe du *Surrey Morphology Group* : Grev et Dunstan, ainsi que Matthew Baerman, Patricia Cabredo-Hofherr, Maris Camilleri, Marina Chumakina, Scott Collier, Penny Everson, Sebastian Fedden, Antonio Fortin, Lisa Mack, Enrique Palancar et Anna Thornton. Merci pour tout ce merveilleux temps passé chez vous, pour votre bonne humeur et pour nos discussions — sans oublier, bien sûr, le café de 11 heures et le chocolat !

Mon séjour à Surrey a été rendu possible grâce à une bourse qui m'a été attribuée par le réseau ESF NetWords. Cette bourse a été pour moi une chance incroyable. Je remercie le comité NetWords de m'avoir attribué la bourse numéro 3774 grâce à laquelle ce travail aura pu murir dans un environnement idéal et qui m'aura permis de rencontrer des personnes exceptionnelles avec lesquelles j'ai eu le plaisir de pouvoir continuer à collaborer depuis. Une partie des données exposées dans cette thèse, en particulier celles du maltais, y figurent par la conséquence directe de collaborations mises en place grâce à ce séjour.

Pour leur soutien logistique et leur gentillesse, je remercie évidemment l'ancien directeur de l'UFR de Linguistique de l'Université Paris Diderot Philippe Martin ainsi le personnel administratif de l'UFRL et du Laboratoire de Linguistique Formelle, en particulier Jeanine, Janique, Denise et Dorothee. Merci aussi à Benoit Chevillon pour ses conseils de dernière minute au moment de l'impression de cette thèse !



---

Je remercie mes collègues de l'Université Paris Sorbonne pour leur accueil à l'UFR de Langue Française.

Pour les rires, les pleurs, les coupes de champagne des pots de mes prédécesseurs, les échanges d'information sur les postes d'ATER à paraître, mais également les discussions stimulantes, je remercie mes co-doctorants et post-doctorants rencontrés à l'UFR de Linguistique et autour, mais notamment au Laboratoire de Linguistique Formelle et à Alpage. Je remercie tout particulièrement Gabriela Bilbàie, Luc Boruta, Jean-Emmanuel Deschaud, Enrique Henestroza, Anna Gazdik, Valérie Hanoka, Pierre Magistry, Elena Ourjoumtseva, Charlotte Roze, Rosa Stern, Jana Strnadová, Sattisvar Tandabany, Elsa Tolone, Yann Strozecki, Grégoire Winterstein. Avec vous j'aurais eu l'occasion de parler de science, de partager des moments de théâtre CIESques, de déplacer bon nombre de cartons, de danser et de boire des cocktails jusque tard dans la nuit dans un vieux bâtiment orange (ou dont la couleur reste encore à qualifier). Enfin un grand merci à Grégoire pour la feuille de style LaTeX dont je suis partie pour la mise en page de cette thèse.

J'aimerais également remercier mes autres collègues, des plus anciens aux plus récents, pour les nombreuses discussions très enrichissantes et non moins agréables qui ont accompagné ces dernières années ou ces derniers mois : Gilles Authier, Marie Candito, Benoit Crabbé, Berthold Crysmann, Bernard Fradin, Romain Garnier, Philippe Gréa, Danièle Godard, Guillaume Jacques, Aimée Lahaussais, Sarah Leroy, Françoise Kerleroux, Pierre Magistry, Jean-Marie Marandin, Philippe Martin, Chris Reintges, Sattisvar Tandabany, Djamá Seddah, Florence Villoing ainsi que tous les autres collègues que j'ai pu rencontrer à différents colloques pour leurs conseils, leurs encouragements et les discussions intéressantes, et dont certains sont devenus des amis. Je pense en particulier à Ash Asudeh, Laurence Danlos, Darja Fišer et Tomaž Erjavec, Matthieu Lafourcade, Enrique Palancar et Marie-Claude Paris. Merci aussi à Nizar Habash et Owen Rambow pour m'avoir invitée au CCLS de l'Université de Columbia pour travailler avec eux dans le cadre de leur projet sur l'analyse automatique de la morphologie : les discussions que nous avons eues m'ont offert un recul précieux sur mon domaine d'étude, la morphologie flexionnelle en général et la morphologie de l'arabe en particulier, et m'ont permis de le redécouvrir par d'autres approches complémentaires et nouvelles pour moi.

Mes étudiants tant à l'Université Paris Diderot qu'à l'Université Paris Sorbonne ont également contribué à la grande richesse de ces dernières années et méritent donc remerciements. Je garde d'excellents souvenirs de ces heures passées en leur compagnie !

J'ai rencontré pendant les années qu'a duré ma thèse un nombre considérable de personnes dont nombreux sont ceux qui, par les discussions que nous avons eues,

---

ont enrichi mon travail et ma conception de la linguistique en général. J'ai essayé de n'oublier personne dans les lignes ci-dessus, mais il est probable qu'au moment d'écrire ces lignes certains noms m'aient échappé. Je prie les personnes concernées de bien vouloir m'en excuser et j'espère qu'elles ne m'en tiendront pas rigueur. Qu'elles soient ici tout particulièrement remerciées pour le rôle qu'elles ont joué et pour cette ultime acte bienveillant.

Enfin, une thèse qui dure plusieurs années ne peut se faire sans le soutien indirect mais non moins essentiel de personnes proches. J'adresse donc ici également un très profond remerciement à mes parents pour tout ce qu'ils ont fait pour moi, pour leur soutien, leur confiance et leur présence à mes côtés même dans les moments les plus durs, ainsi qu'à ma toute famille. Je suis également fier de pouvoir remercier mes amis, fidèles pendant tout ce temps et sur qui j'ai appris que je pouvais compter, tout particulièrement David, mais aussi Elena, Heike, Magali, Olivier, Pierre, Satti et Yann. Quant à Benoît, au delà de ce qui a déjà été dit, je le remercie aussi pour tout le reste : pendant les longs mois de rédaction, sa présence indiscutable et son soutien infaillible n'auront eu d'égal que sa patience.

Paris, le 2 juin 2013  
Géraldine Walther



# Résumé de la thèse

L'objectif principal de cette thèse est de proposer un modèle formel de la morphologie flexionnelle, nommé  $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$  (*Paradigm Shape and Lexicon Interface*) qui soit typologiquement fondé et qui permette de concevoir, d'implémenter et d'évaluer quantitativement des analyses formelles de systèmes flexionnels particuliers. Il repose pour cela sur les concepts de la typologie canonique (Corbett, 2003, 2007a). Fondé sur une approche inférentielle-réalisationnelle de la morphologie (Stump, 2001), ce modèle permet notamment de formuler explicitement le rapport qu'entretient le lexique avec les caractéristiques morphologiques observables dans les paradigmes flexionnels des lexèmes et d'évaluer la part de régularité et d'irrégularité flexionnelle inhérente à la définition de chacun d'entre eux.

$\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$  est également implémenté dans un formalisme dédié, *Alexina* $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$  (Sagot & Walther, 2013), qui constitue une nouvelle version du formalisme lexical *Alexina* (Sagot, 2010). *Alexina* est à l'origine un formalisme lexical conçu pour le traitement automatique des langues (TAL) et dans lequel il existe un nombre important de ressources lexicales libres à grande et moyenne échelle pour des langues diverses.

Enfin,  $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$  s'accompagne d'un éventail de mesures quantitatives. La première est une *mesure de compacité* reposant sur le concept de longueur de description minimale (Minimal Description Length ou MDL (Rissanen, 1984)) empruntée à la théorie de l'information (Walther & Sagot, 2011). Elle permet d'évaluer l'économie descriptive d'une description donnée et de la comparer quantitativement à des descriptions concurrentes. Dans cette thèse, cette mesure permet notamment d'évaluer quantitativement la compacité de quatre descriptions implémentées concurrentes de la flexion verbale du français (env. 7 800 verbes, 370 000 formes), de deux descriptions concurrentes de la flexion verbale du latin (env. 2 300 verbes, 260 000 formes), ainsi que de deux descriptions implémentées d'environ 600 verbes (9 000 formes) du maltais.

Le second ensemble de mesures est un ensemble de mesures de canonicité permettant

---

de quantifier le degré de canonicité observable dans les paradigmes isolés ou les systèmes flexionnels complets d'une langue particulière décrite en פֿאַרשטאַנד. Il s'agit ainsi d'un moyen de rendre quantifiable la notion qualitative de canonicité flexionnelle développée par Corbett (2007a). Dans cette thèse, cet ensemble de mesures a été appliqué (en plus de la mesure de compacité) pour comparer deux descriptions du système verbal latin (env. 2 300 verbes, 260 000 formes), dont l'une défend une approche dérivationnelle du passif latin tandis que l'autre suit l'analyse flexionnelle plus traditionnelle et a montré l'influence des choix descriptifs sur les caractéristiques typologiques (canoniques) des systèmes décrits.

En parallèle, ma thèse aborde également, à partir de descriptions de systèmes flexionnels, les questions de la frontière entre la morphologie et les autres composantes de la langue, notamment par l'étude de données du maltais (interface morphologie flexionnelle/phonologie), du latin (interface morphologie flexionnelle/dérivationnelle) et du kurde sorani (interface morphologie flexionnelle/syntaxe).

# Abstract of the thesis

The main aim of this thesis consists in the introduction of a new formal model of inflectional morphology called  $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}\mathcal{I}$  (*Paradigm Shape and Lexicon Interface*). It is based on typological theory and allows for conceiving, implementing, and quantitatively evaluating formal analyses of specific inflectional systems. It relies on concepts established within the framework of Canonical Typology developed by Greville G. Corbett (2003; 2007a). Based on an inferential-realisation approach to morphology (Stump, 2001), this model in particular allows for explicitly formulating the relationship between the lexicon and the morphological properties that can be observed within the inflectional paradigms of individual lexemes, and for evaluating the regularity or irregularity of each one of these properties.

$\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}\mathcal{I}$  has also been equipped with its own implementation tools called *Alexina $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}\mathcal{I}$*  (Sagot & Walther, 2013), which constitute a newly adapted version of the *Alexina* lexical framework (Sagot, 2010). *Alexina* itself is a framework originally developed in order to serve as a lexical framework for building large-scale lexical resources, mainly to be used within various types of NLP tools and for which a significant number of large and medium scale resources exists.

Finally  $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}\mathcal{I}$  comprises a set of quantitative measures of language structure and language description evaluation. The first measure is an *descriptive economy* or *compactness measure* which relies on the information theoretical concept of *Minimal Description Length* (MDL) Rissanen (1984). This measure allows for assessing the relative descriptive economy of competing descriptions for the same set of data (encoded in *Alexina $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}\mathcal{I}$* ). Within this thesis, this measure has been used for contrastively evaluating the descriptive of four implemented descriptions for French verbs (approx. 7,800 verbs, 370,000 forms), two descriptions of Latin verbal inflection (approx. 2,300 verbs, 260,000 forms), and two competing descriptions of approximately 600 Maltese verbs (approx. 9,000 forms).

---

The second set of measures is a set of *canonicity measures* allowing for quantifying the degree of canonicity displayed by a given inflectional paradigm or even, globally, an whole inflectional system. The aim in developing this set of measures was to provide the traditional qualitative canonical approach with specifically designed measures for quantifying degrees of deviation from the canon. It has been applied to the two above mentioned competing descriptions of Latin verbal inflection.

In addition to the development of  $\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}$ , this thesis also addresses the question of interfaces of inflectional morphology with other components of grammar or language. Examples developed within the thesis concern the morphology/lexicon interface in French, morphology/phonology interface in Maltese first binyan verbs, the inflectional *versus* derivational morphology interface in Latin and the morphology/syntax interface in Sorani Kurdish.

# Notations dans PARSL



1. Un système morphologique flexionnel :  $\mu_{\text{infl}}$
2. Les traits morphosyntaxiques :
  - Un ensemble de paquets de traits morphosyntaxiques (PTM) :  $\Sigma$
  - Un paquet de traits morphosyntaxiques particulier :  $\sigma$
3. Le lexique :
  - Une catégorie flexionnelle :  $\mathcal{C}$
  - Un ensemble de catégories flexionnelles :  $\mathcal{C}$
  - Le lexique flexionnel :  $\mathcal{L}$
  - Une entrée lexicale :  $i$
  - La base phonologique donnée en entrée aux règles de flexion :  $b$
  - Un ensemble de radicaux supplétifs :  $S_i^{\text{suppl}}$
  - Un ensemble de formes supplétives :  $F_i^{\text{suppl}}$
  - Un schème flexionnel :  $\mathbb{P}$
  - Un sous-schème :  $\mathbb{P}$
  - Un couple réalisationnel :  $p = (z, t)$
  - Une zone réalisationnelle :  $z$
  - Une fonction de transfert :  $t$
4. La grammaire
  - Un espace partitionnant :  $\pi$
  - Une zone/classe de radicaux :  $z^s/Z^s$
  - Une zone/classe thématique :  $z^h/Z^h$
  - Une zone/classe d'exponence/flexionnelle :  $z^{\text{exp}}/Z^{\text{exp}}$
  - Une zone/classe réalisationnelle :  $z/Z$
  - Un niveau réalisationnel :  $l$
  - Un ensemble de  $n$  niveaux réalisationnels :  $\{l_1, \dots, l_n\}$
  - Une règle de sélection de couples réalisationnels :  $\mathbb{L}_l$
  - Un ensemble de règles de sélection de couples réalisationnels :  $\{\mathbb{L}_{l_1}, \dots, \mathbb{L}_{l_n}\}$
  - Une règle de sélection de zones réalisationnelles :  $\mathbb{L}_l|Z$
  - Un ensemble de règles de sélection de zones réalisationnelles :  $\{\mathbb{L}_{l_1}|Z, \dots, \mathbb{L}_{l_n}|Z\}$
  - Une règle de réalisation :  $R$
  - Le résultat d'une règle de réalisation :  $r$
5. Les objets construits
  - Un radical :  $s$
  - Une forme :  $f$
  - Un paradigme :  $P$
  - Une case dans un paradigme :  $C$

FIGURE 0.1 : Notation des éléments de la grammaire dans  $\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}$

---

LÉGENDE :

$i$	une entrée lexicale dans un lexique flexionnel $\mathcal{I}$	
I-PHON	une séquence phonologique originale servant d'entrée aux règles de réalisation	
I-CAT	la catégorie flexionnelle de $i$ notée $\mathcal{C}_i$ appartenant à l'ensemble $\mathcal{C}$ des catégories flexionnelles d'une langue	
MSF	un ensemble de traits morphosyntaxiques exprimés par $i$ et noté $\Sigma_i$	
$\sigma$	un paquet de traits morphosyntaxiques (PTM)	
S-STEM	un ensemble canoniquement vide de radicaux supplétifs, un radical noté $s$ étant constitué de sa zone de radical $z^s$ et de sa réalisation formelle $r^s$	
S-FORM	un ensemble canoniquement vide de formes supplétives, une forme $f$ étant constituée du PTM $\sigma$ qu'elle exprime et de sa réalisation formelle $r^f$	
I-PAT	le schème flexionnel de $i$ comprenant un ensemble de sous-schémas, noté $\mathbb{P}_i$	
$\mathbb{P}$	un sous-schème flexionnel comportant des couples réalisationnels notés $p$ de forme $(z, t)$	
$z$	une zone réalisationnelle	
$t_z$	la fonction de transfert associée à $z$ dans un couple réalisationnel $p$	

$i$	<table border="0" style="width: 100%;"> <tr> <td style="vertical-align: top;">I-PHON</td> <td><math>b_i</math></td> </tr> <tr> <td style="vertical-align: top;">I-CAT</td> <td><math>\mathcal{C}_i</math></td> </tr> <tr> <td style="vertical-align: top;">MSF</td> <td><math>\Sigma_i = \{\sigma_1, \dots, \sigma_n\}</math></td> </tr> <tr> <td style="vertical-align: top;">S-STEM</td> <td><math>\{(z_1^s, r_1^s), \dots, (z_n^s, r_n^s)\}</math></td> </tr> <tr> <td style="vertical-align: top;">S-FORM</td> <td><math>\{(\sigma_1, r_1^f), \dots, (\sigma_n, r_n^f)\}</math></td> </tr> <tr> <td style="vertical-align: top;">I-PAT</td> <td> <math>\mathbb{P}_1 = p_{1,1}, \dots, p_{1,m}</math>  <math>\vdots</math>  <math>\mathbb{P}_n = p_{n,1}, \dots, p_{n,m}</math> </td> </tr> </table>	I-PHON	$b_i$	I-CAT	$\mathcal{C}_i$	MSF	$\Sigma_i = \{\sigma_1, \dots, \sigma_n\}$	S-STEM	$\{(z_1^s, r_1^s), \dots, (z_n^s, r_n^s)\}$	S-FORM	$\{(\sigma_1, r_1^f), \dots, (\sigma_n, r_n^f)\}$	I-PAT	$\mathbb{P}_1 = p_{1,1}, \dots, p_{1,m}$ $\vdots$ $\mathbb{P}_n = p_{n,1}, \dots, p_{n,m}$
I-PHON	$b_i$												
I-CAT	$\mathcal{C}_i$												
MSF	$\Sigma_i = \{\sigma_1, \dots, \sigma_n\}$												
S-STEM	$\{(z_1^s, r_1^s), \dots, (z_n^s, r_n^s)\}$												
S-FORM	$\{(\sigma_1, r_1^f), \dots, (\sigma_n, r_n^f)\}$												
I-PAT	$\mathbb{P}_1 = p_{1,1}, \dots, p_{1,m}$ $\vdots$ $\mathbb{P}_n = p_{n,1}, \dots, p_{n,m}$												

---

FIGURE 0.2 : Notation pour une entrée lexicale standard dans PARSLI.



# Table des matières

Remerciements	i
Résumés	vi
Résumé de la thèse	vi
Abstract of the thesis	ix
Notations dans $\mathcal{PARSL}$	xi
Introduction	7
Introduction	11
I État de l’art pour une approche typologique et formelle de la morphologie flexionnelle	27
1 Méthodologie pour l’analyse formelle en morphologie : le cas des déponents latins	29
1.1 Présentation générale du passif en latin . . . . .	32
1.1.1 La flexion verbale du latin . . . . .	32

1.1.2	Le passif en latin . . . . .	42
1.2	Deux analyses contemporaines des déponents latins . . . . .	50
1.2.1	Analyse flexionnelle des déponents latins : Hippiusley (2007) . . . . .	51
1.2.2	Une approche lexicale des déponents latins : Kiparsky (2005) . . . . .	60
1.3	Les « déponents latins », des paradigmes hétéroclites . . . . .	65
1.3.1	Le passif latin : dérivation ou flexion ? . . . . .	66
1.3.2	Expliquer les paradigmes des verbes déponents . . . . .	68
1.4	Conception d'une nouvelle analyse formelle de la flexion latine . . . . .	71
1.4.1	Représentation de l'hétéroclise . . . . .	71
1.4.2	Représentation des niveaux de réalisation des formes . . . . .	76
1.4.3	Classes flexionnelles des verbes athématiques et à thème en <i>i</i> . . . . .	82
1.4.4	Modélisation des règles de réalisation . . . . .	82
1.4.5	La défektivité des verbes latins . . . . .	90
1.4.6	Extension formelle de l'analyse . . . . .	91
1.5	Avantages et problèmes liés à la formalisation de la morphologie . . . . .	93
<b>2</b>	<b>Approches formelles de la morphologie flexionnelle</b>	<b>95</b>
2.1	Délimitation du cadre conceptuel . . . . .	96
2.1.1	Définition du lexicalisme . . . . .	96
2.1.2	La morphologie autonome . . . . .	101
2.1.3	Objet d'étude de la morphologie . . . . .	102
2.1.4	Une vision réalisationnelle de la morphologie . . . . .	104
2.1.5	Flexion et construction et ce qui les distingue . . . . .	107
2.1.6	La notion de paradigme . . . . .	111
2.1.7	Traitement de l'allomorphie de radical . . . . .	112
2.1.8	La structure du lexique . . . . .	114
2.2	Conception modulaire articulée de la grammaire . . . . .	121
2.3	Approches inférentielles-réalisationnelles . . . . .	125
2.3.1	<i>Paradigm Function Morphology</i> Stump (2001) . . . . .	126
2.3.2	La théorie flexionnelle en PFM . . . . .	131
2.3.3	<i>Network Morphology</i> Brown & Hippiusley (2012) . . . . .	136
<b>3</b>	<b>Qualifier la différence : la typologie canonique</b>	<b>145</b>
3.1	Définir les concepts fondamentaux . . . . .	148
3.1.1	Un exemple de définition conceptuelle : la réduplication . . . . .	148
3.1.2	Résumé des propriétés de définitions notionnelles pour la typologie	151

3.2	Fondamentaux et prototypes . . . . .	151
3.2.1	Définir des universaux linguistiques . . . . .	153
3.2.2	La notion de prototype . . . . .	155
3.3	La morphologie naturelle . . . . .	156
3.3.1	Motivations pour une morphologie naturelle . . . . .	156
3.3.2	Paramètres de la morphologie naturelle . . . . .	157
3.3.3	Naturalité globale et naturalité locale . . . . .	159
3.3.4	Naturalité ou autonomie . . . . .	160
3.4	Fondamentaux et canons . . . . .	161
3.4.1	Principes généraux de la typologie canonique . . . . .	161
3.4.2	Ontologies et canons . . . . .	164
3.4.3	La morphologie canonique . . . . .	165
3.4.4	La flexion canonique . . . . .	167
3.5	Phénomènes flexionnels canoniques et non canoniques . . . . .	171
3.5.1	Non-canonlicité de la forme . . . . .	172
3.5.2	Non-canonlicité du remplissage des cases des paradigmes . . . . .	178
3.5.3	Non-canonlicité dans l'expression des traits morphosyntaxiques . . . . .	181
3.5.4	Huit phénomènes non canoniques : résumé . . . . .	184
3.6	Prototypes, formes naturelles et canon en typologie . . . . .	185

## II Formaliser une approche typologique de la morphologie flexionnelle 191

4	$\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}\mathcal{I}$ , un modèle formel de la morphologie flexionnelle <span style="float: right;">193</span>
4.1	Description des propriétés spécifiques de $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}\mathcal{I}$ . . . . . <span style="float: right;">195</span>
4.1.1	Zones flexionnelles . . . . . <span style="float: right;">195</span>
4.1.2	Représentation de l'hétéroclise . . . . . <span style="float: right;">198</span>
4.1.3	Classes de radicaux . . . . . <span style="float: right;">199</span>
4.1.4	Le niveau thématique . . . . . <span style="float: right;">204</span>
4.1.5	Une représentation multi-niveaux . . . . . <span style="float: right;">205</span>
4.1.6	Représentation des phénomènes de décalage morphosyntaxique . <span style="float: right;">214</span>
4.1.7	Représentation de la surabondance, de la défektivité et de la surdifférenciation . . . . . <span style="float: right;">219</span>

4.1.8	Résumé : modélisation de dix phénomènes non canoniques en $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}\mathcal{I}$ . . . . .	224
4.2	Présentation des primitives du modèle . . . . .	228
4.2.1	Entrées lexicales flexionnelles . . . . .	228
4.2.2	Les traits morphosyntaxiques . . . . .	232
4.2.3	Les catégories flexionnelles . . . . .	232
4.2.4	Les schèmes flexionnels . . . . .	235
4.2.5	Construction des formes et phénomènes de supplétion . . . . .	246
4.2.6	Les classes flexionnelles et les classes de radicaux . . . . .	247
4.3	$\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}\mathcal{I}$ , PFM et <i>Network Morphology</i> . . . . .	247
4.4	Quantifier la différence : les mesures de canonicité dans $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}\mathcal{I}$ . . . . .	249
4.4.1	Mesures du remplissage des cases du paradigme . . . . .	250
4.4.2	Mesures de la canonicité des réalisations . . . . .	262
4.4.3	Mesures de réalisation canoniques des traits morphosyntaxiques . . . . .	267
4.5	Résumé et discussion . . . . .	274
<b>5</b>	<b>Implémentation de descriptions</b> . . . . .	<b>277</b>
5.1	Pourquoi implémenter ? . . . . .	277
5.1.1	Approches théoriques et implémentations . . . . .	278
5.1.2	Les approches computationnelles . . . . .	278
5.1.3	De l'importance de la grande échelle . . . . .	280
5.2	Implémenter $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}\mathcal{I}$ . . . . .	282
5.2.1	Le formalisme Alexina originel . . . . .	283
5.2.2	Le formalisme Alexina adapté à $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}\mathcal{I}$ . . . . .	290
5.3	Implémentation d'une description du système verbal du maltais . . . . .	313
5.3.1	La flexion verbale maltaise . . . . .	313
5.3.2	Analyse des verbes maltais selon Camilleri (2011) . . . . .	319
5.3.3	Analyse des formes complètes dans $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}\mathcal{I}$ . . . . .	329
5.4	Analyses implémentées et choix descriptifs . . . . .	338
<b>6</b>	<b>Évaluations externes de descriptions morphologiques</b> . . . . .	<b>343</b>
6.1	Mesurer la longueur de description . . . . .	349
6.1.1	La complexité de Kolmogoroff . . . . .	350
6.1.2	La longueur de description minimale (LDM) . . . . .	351
6.2	Répartition de l'information : lexicale ou grammaticale . . . . .	354
6.2.1	Représenter la flexion verbale du français . . . . .	354

6.2.2	Mesurer la complexité de différentes descriptions morphologiques de la flexion verbale du français . . . . .	358
6.2.3	Compacité et formalisation . . . . .	362
6.3	Filtrer l'information : morphologie vs. phonologie . . . . .	363
6.3.1	Analyse phonologique des radicaux maltais . . . . .	364
6.3.2	Mesures de longueurs de description des deux analyses . . . . .	387
6.3.3	L'interface phonologie-morphologie et les cas de morphologie autonome . . . . .	389
6.4	Approche syntaxique ou morphologique ? . . . . .	391
6.4.1	La flexion verbale et le marquage personnel en kurde sorani . . . . .	392
6.4.2	Analyse du placement des MPM pour les formes n'exprimant pas le parfait . . . . .	397
6.4.3	Prise en compte des formes du parfait . . . . .	402
6.4.4	Conclusion sur la structure à l'interface de la morphologie et de la syntaxe . . . . .	406
6.5	Composantes de la grammaire et répartition de l'information . . . . .	407
<b>7</b>	<b>Mesures de canonicité et longueur de description</b>	<b>409</b>
7.1	Formalisation de la description de la flexion des verbes latins en $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$ . . . . .	410
7.1.1	Définition de la catégorie flexionnelle . . . . .	410
7.1.2	Définition des trois niveaux réalisationnels principaux . . . . .	416
7.1.3	Combinaison inter-niveaux des classes réalisationnelles . . . . .	424
7.1.4	Représentation de la dérivation passive . . . . .	425
7.1.5	Représentation de verbes déponents . . . . .	426
7.1.6	Les schèmes flexionnels de l'analyse $3\mathcal{H}$ . . . . .	428
7.2	Proposition d'une description à quatre niveaux de la flexion des verbes latins	428
7.2.1	Les classes réalisationnelles définies dans l'analyse $4\mathcal{H}$ . . . . .	429
7.2.2	Interaction des classes réalisationnelles à travers les niveaux dans l'analyse $4\mathcal{H}$ . . . . .	432
7.2.3	Représentation de la dérivation passive dans l'analyse $4\mathcal{H}$ . . . . .	433
7.2.4	Modélisation des verbes déponents dans l'analyse $4\mathcal{H}$ . . . . .	434
7.3	Mesure de la longueur de description des analyses concurrentes . . . . .	435
7.3.1	Comparaison de la longueur de description des analyses $3\mathcal{H}$ et $4\mathcal{H}$ . . . . .	435
7.3.2	Comparaison de la longueur de description de nos analyses avec l'analyse traditionnelle . . . . .	442



7.4	Mesures de canonicité . . . . .	448
7.4.1	Identification des phénomènes non canoniques présents . . . . .	449
7.4.2	Mesures de la réalisation canonique des traits morphosyntaxiques	453
7.4.3	Mesures du remplissage des cases des paradigmes . . . . .	458
7.4.4	Mesures de la canonicité des réalisations . . . . .	464
7.5	Discussion et perspectives . . . . .	472
 <b>Conclusion générale</b>		<b>476</b>
 <b>Annexes</b>		<b>487</b>
<b>A</b>	<b>Implémentations des descriptions morphologiques</b>	<b>489</b>
A.1	Le système verbal du latin . . . . .	489
A.1.1	Formalisation des règles de réalisation en <i>Paradigm Function Morphology</i> . . . . .	489
A.1.2	Implémentation native en Alexina . . . . .	492
A.1.3	Implémentation de la variante de l'analyse en trois niveaux avec hétéroclise radicale en Alexina . . . . .	493
A.1.4	Implémentation de l'analyse <i>Network Morphology</i> en DATR . . . . .	493
A.1.5	Implémentation de l'analyse à quatre niveaux du système verbal latin en Alexina <sub>PARSLI</sub> . . . . .	493
A.1.6	Implémentation de la variante de l'analyse en quatre niveaux avec hétéroclise radicale en Alexina . . . . .	494
A.1.7	Implémentation de l'analyse traditionnelle en trois niveaux en Alexina . . . . .	494
A.2	Le premier binyan du système verbal à base sémitique du maltais . . . . .	494
A.2.1	Implémentation de l'analyse morphologique de la flexion du premier binyan en Alexina <sub>PARSLI</sub> . . . . .	494
A.2.2	Implémentation de l'analyse morphologique de la flexion du premier binyan en Alexina <sub>PARSLI</sub> . . . . .	495
<b>B</b>	<b>Schémes flexionnels distincts dans Leffla</b>	<b>497</b>
B.1	Analyses 3H et 4H . . . . .	497
B.2	Analyse 3F . . . . .	499

<b>C</b>	<b>Interaction avec une théorie indépendante de la notion de traits morphosyntaxiques</b>	<b>501</b>
C.1	Une théorie canonique des hiérarchies de traits . . . . .	501
C.1.1	Sous-spécification et neutralisation . . . . .	501
C.1.2	Surdifférenciation de valeurs et de traits . . . . .	502
C.1.3	Difficultés liées à l'interprétation de la non-canonieité dans les hiérarchies de traits . . . . .	503
C.2	Représentation formelle de la neutralisation et de la surdifférenciation de traits . . . . .	505
C.3	Alignement des formes et des hiérarchies . . . . .	506
C.4	Morphomes et hiérarchies de traits . . . . .	507
C.5	Encodage de la morphomicité dans $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$ . . . . .	508
<b>D</b>	<b>Transcription du sorani</b>	<b>511</b>
	<b>Index</b>	<b>515</b>
	<b>Bibliographie</b>	<b>529</b>
	<b>Bibliographie</b>	<b>529</b>



# Introduction



# Introduction

L'objectif de cette thèse est de proposer un modèle formel de la morphologie flexionnelle qui soit typologiquement fondé et qui permette de concevoir, d'implémenter et d'évaluer quantitativement des analyses formelles de systèmes flexionnels particuliers.

## Typologie et morphologie formelle

L'étude de la morphologie flexionnelle se heurte à un paradoxe apparent : les données qu'elle traite sont fondamentalement faciles à décrire, mais c'est justement cette facilité à les décrire qui fait apparaître clairement l'extrême divergence entre les différents systèmes flexionnels et qui révèle ainsi toute la complexité du domaine.

Si l'on se situe dans une perspective synchronique stricte, les données traitées en morphologie flexionnelle sont finies : la systématisme de la morphologie flexionnelle fait que le nombre de formes dont elle doit rendre compte ne sera jamais plus qu'un multiple du nombre d'entrées lexicales contenues à un instant donné dans le lexique de la langue<sup>1</sup>. Le nombre de règles de réalisation de ces formes sera donc également fini, de sorte qu'il est aisé de décrire intégralement un système flexionnel. Le morphologue flexionnel bénéficie en cela d'un net avantage sur le syntacticien ou même sur le morphologue constructionnel qui traitent à chaque moment d'un nombre non fini ou pour le moins très vaste de données. Par ailleurs, les données traitées par la syntaxe et la morphologie constructionnelle sont le plus souvent soumises à des problèmes d'acceptabilité, là où la morphologie flexionnelle permet généralement d'identifier les règles construisant les formes attestées et qui sont donc

---

1. Il va de soi que les procédés de création de lexèmes nouveaux modifient à chaque instant l'étendue du lexique d'une langue donnée. Cependant on ne crée pas de cette manière d'irrégularités isolées, si bien que la description morphologique peut toujours prendre la forme d'une liste finie de règles accompagnée d'une liste finie d'irrégularités.

---

à décrire<sup>2</sup>. Une description comme celle du Bescherelle (Arrivé, 1997) apparaît ainsi comme une description empiriquement satisfaisante de la flexion verbale du français, même si l'on peut la considérer comme théoriquement inintéressante.

En conséquence, l'enjeu de l'étude de la morphologie flexionnelle ne peut se limiter à la simple description des systèmes flexionnels existants. Les questions ouvertes concernent essentiellement deux domaines : l'évaluation des descriptions proposées et, surtout, la capacité à rendre compte de façon satisfaisante et homogène de la diversité des systèmes flexionnels à travers les langues.

Le second de ces enjeux est en fait celui de la TYPLOGIE LINGUISTIQUE. Ce sont les approches typologiques qui permettent, par des études empiriques, d'appréhender la richesse des données et de mettre en place un VOCABULAIRE DESCRIPTIF EMPIRIQUEMENT FONDÉ.

Considérons par exemple le biais qu'ont pu introduire les études sur les langues romanes dans la définition de la notion de MORPHEME. Cette notion apparaît pour la première fois en tant que *niveau morphomique* dans (Aronoff, 1994). Aronoff illustre cette notion en notant le fait que les alternances de radicaux en latin ne coïncident pas avec des ensembles cohérents de traits morphosyntaxiques<sup>3</sup>. Elle a ensuite été reprise dans les études portant sur la flexion verbale dans les langues romanes (Maiden, 1992 ; Pirrelli & Battista, 2000 ; Bonami & Boyé, 2002, 2003), qui, sans surprise, présentaient des comportements similaires au latin et exhibaient à chaque fois un type d'alternance de radicaux non motivée. Une retombée de ces études est l'introduction des concepts de PARTITIONS de Pirrelli & Battista (2000) et d'ESPACE THÉMATIQUE de Bonami & Boyé (2002, 2003) qui formalisent l'allomorphie radicale en partitionnant les paradigmes en zones utilisant uniformément le même radical. Le fait que les études mettant en jeu la notion de morphème aient systématiquement eu pour objet la flexion des langues romanes a pu donner l'impression que la notion de morphème concerne exclusivement l'alternance radicale. Mais les langues romanes à elles seules constituent un ensemble de langues peu représentatif d'un point de vue typologique, notamment pour établir la définition d'une notion théorique. En particulier, une question qui n'a pas été posée explicitement à cette époque est celle de savoir si les structures morphomiques sont effectivement restreintes au

---

2. À la variation sociolinguistique près.

3. En l'occurrence, chaque verbe est défini pour trois radicaux : un RADICAL 1 pour toutes les formes finies imperfectives, ainsi que celle de l'infinitif présent, du participe présent et du gérondif ; un RADICAL 2 pour les formes finies perfectives (actives) et l'infinitif passé ; un RADICAL 3 employé pour l'infinitif et le participe futurs ainsi que pour le supin.

---

domaine radical ou si on en retrouve également dans le système exponentiel.

Or, comme l'ont ultérieurement montré eux-mêmes Bonami & Boyé (2010b), on trouve également des partitions structurant les systèmes désinentiels. Les données du tableau 0.1 illustrent sur un cas particulier la manière dont les sous-paradigmes de tous les lexèmes verbaux du népal<sup>4</sup> sont structurés par un certain nombre de syncrétismes dont trois sont systématiques (cf. les différentes teintes indiquées dans les deux tableaux). D'autres comme l'alternance entre les marques *-aj-* et *-i-* sont réduites à certains sous-paradigmes les comportant. Ces syncrétismes portent sur des cases qui, comme dans le cas des morphèmes des langues romanes, ne coïncident pas avec des alternances de traits morphosyntaxiques. Les désinences *-aj-* (en bleu) et *-i-* (en vert) du sous-paradigme du présent négatif expriment ainsi respectivement les traits suivants. Au présent négatif,

- la marque *-i-* est employée pour la première personne du singulier et toutes les formes du féminin ;
- la marque *-aj-* est employée pour toutes les formes du masculin hormis la première personne, ainsi que pour l'intégralité du pluriel.

On reconnaît là des distributions morphomiques au sens de la tradition issue d'Aronoff (1994). Elles se retrouvent toutes à travers les quatre classes flexionnelles habituellement posées pour le népal, et certaines (marquées par les cases colorées dans le tableau 0.1) même à travers tous les sous-paradigmes. Une approche typologiquement fondée consistant à observer les zones d'interprétabilité dans les paradigmes de langues plus diverses permet ainsi de montrer la pertinence générale de la notion théorique de morphème, qui n'est pas réduite à la partie radicale des paradigmes. Le modèle de la morphologie flexionnelle que nous introduisons dans cette thèse aura ainsi entre autres la caractéristique de généraliser le systèmes des partitions à la Pirrelli & Battista (2000) ou des espaces thématiques à la Bonami & Boyé (2003) au système exponentiel.

En plus de constituer un fondement indispensable à la définition des concepts préalables à la conception d'analyses morphologiques, ce sont aussi les approches typologiques qui permettent d'évaluer les résultats des analyses flexionnelles en les replaçant dans le tableau plus général des systèmes flexionnels dans les langues du monde.

Notre travail dans cette thèse s'intéresse de près au rôle de la formalisation et de l'implémentation dans la conception d'analyses de systèmes morphologiques flexionnels. Nous considérons la formalisation comme un moyen d'accéder à des généralisations

---

4. Les sous-paradigmes représentés ici sont ceux du passé simple et du présent négatif, mais les auteurs indiquent que la même structure se retrouve dans les autres sous-paradigmes, quels que soient les verbes.



BIRSANU « oublier »						
	M.SG	F.SG	PL	M.SG	F.SG	PL
1	birsē	birsē	birsjaũ	birsād-i-nã	birsād-i-nã	birsād-aj-naũ
2.LOW	birsis	birsis	birsjau	birsād-aj-nas	birsād-i-nas	birsād-aj-nau
2.MID	birsjau	birsjau	birsjau	birsād-aj-nau	birsād-i-nau	birsād-aj-nau
3.LOW	birsjo	birsi	birse	birsād-aj-na	birsād-i-na	birsād-aj-nan
3.MID	birse	birsin	birse	birsād-aj-nan	birsād-i-nan	birsād-aj-nan
	PASSÉ SIMPLE			PRÉSENT NÉGATIF		

TABLE 0.1 : Flexion verbale du népali (données de Bonami & Boyé (2010b))

sur la langue en isolant des questions générales posées par les données particulières. Le but d'une formalisation n'est pas de se substituer à l'étude empirique, mais d'apporter une explicitation formelle et des outils de quantification à des approches typologiques existantes. En cela, une formalisation prend toute sa force si elle est accompagnée de moyens d'implémentation.

Notre thèse s'inscrit ainsi dans la tradition de la linguistique formelle, qui préconise une FORMALISATION explicite pour garantir LA COHÉRENCE DANS L'EMPLOI DES CONCEPTS définis sur des bases empiriques. En ce sens, elle s'inscrit dans la lignée des approches descendant de la tradition des fragments de grammaires initiée par Montague (1970). Nous défendons cependant la thèse selon laquelle une formalisation des données ne tire toute sa puissance pour le traitement de questions théoriques posées qu'en implémentant des fragments larges des phénomènes étudiés. Dans le cas de la morphologie flexionnelle, ces fragments peuvent être quasi-exhaustifs, c'est-à-dire prendre en compte une portion du lexique de la langue proche de sa totalité<sup>5</sup>. Nous allons donc au delà de la méthode du fragment et nous nous appuyons sur des IMPLÉMENTATIONS À GRANDE ÉCHELLE qui assurent à la fois LA COMPLÉTUDE ET LA COHÉRENCE GLOBALE, c'est-à-dire avec le système de la langue dans son ensemble. En donnant accès à une quantité réaliste de données, l'implémentation à grande échelle permet de ne pas omettre des cas particuliers, et surtout de ne pas surévaluer l'importance de phénomènes marginaux ne concernant que quelques cas isolés par rapport au fonctionnement général du système linguistique étudié. Nous développerons ainsi dans cette thèse les arguments justifiant l'implémentation systématique et à grande échelle de descriptions morphologiques flexionnelles.

5. Les frontières du lexique étant instables et fluctuantes, l'exhaustivité stricte est évidemment hors de portée. La quasi-exhaustivité est cependant suffisante pour avoir une vision réaliste de l'ampleur des phénomènes.

---

Une fois les concepts pertinents définis, la modélisation d'une analyse cohérente d'un système flexionnel donné devient toutefois en soi une tâche relativement simple. Il est ainsi souvent possible de rendre compte d'un système flexionnel donné par plusieurs analyses concurrentes, également cohérentes et difficiles à départager sur des bases conceptuelles. Par exemple, pour la conjugaison du français, des analyses aussi différentes que celles de Plénat (1987), Swiggers & van den Eynde (1987), Paradis & Fenne (1995), Bonami & Boyé (2003) et Kilani-Schoch & Dressler (2005) reposent sur les mêmes outils analytiques de base : une segmentation en radical et exposants, la possibilité d'allomorphies radicales, l'organisation possible des exposants en classes flexionnelles et des règles morphologiques. Ce ne sont donc pas les primitives analytiques mais leur emploi à des degrés divers qui produit la différence entre ces analyses. Quel que soit le système flexionnel étudié, il sera ainsi toujours possible d'en rendre compte par plusieurs analyses pour certaines très différentes, même à l'intérieur d'un cadre descriptif donné.

Une manière dont certains auteurs ont proposé de mettre un terme au PROBLÈME DES ANALYSES CONCURRENTES consiste à limiter l'utilisation d'un certain nombre de concepts théoriques en adoptant une APPROCHE ABSTRACTIVE plus directe, voir notamment (Blevins, 2003, 2006 ; Ackerman et al., 2009 ; Blevins, 2012)<sup>6</sup>. Contrairement AUX APPROCHES CONSTRUCTIVES précédemment citées, ces dernières tentent de tirer des généralisations sur les systèmes flexionnels à partir des formes brutes. Elles ne dépendent notamment pas de décisions arbitraires quant à la segmentation des formes, qui font notamment la différence majeure entre les descriptions de Plénat (1987), Swiggers & van den Eynde (1987), Paradis & Fenne (1995), Bonami & Boyé (2003) et Kilani-Schoch & Dressler (2005) évoquées ci-dessus. Mais ces approches ont en retour le défaut de se priver en l'état actuel des choses d'une partie importante de l'éclairage conceptuel mis en place par les approches typologiques.

Dans la mesure où nous défendons le point de vue que l'objectif de l'analyse flexionnelle est à terme de pouvoir interpréter dans une démarche typologique les particularités d'un système flexionnel à la lumière d'autres systèmes, cette solution ne saurait en être une. Dans cette thèse, nous tentons donc de répondre différemment à LA QUESTION DE L'ARBITRAIRE sous-tendant la multiplicité des analyses. Plutôt que de l'évacuer par une approche abstraite, nous proposons d'*accepter la part d'arbitraire que comportent*

---

6. Les approches abstraites reposent sur des principes d'analogie entre cases de paradigmes plutôt que sur la recherche de règles de réalisation partielles permettant de décrire la construction des formes segmentées à partir de généralisations partielles sur des radicaux ou des exposants définis au préalable (Blevins, 2012). La question première dans l'étude de la morphologie flexionnelle devient alors la question abstraite du degré d'incertitude sur le mode de remplissage des cases des paradigmes, le *Paradigm Cell Filling Problem (PCFP)* (Ackerman et al., 2009)

---

*intrinsèquement les approches constructives* et de chercher un moyen de la contenir par des évaluations objectives externes. Le critère que nous avons choisi dans cette thèse est un critère objectif possible parmi d'autres. Il s'appuie toutefois sur une idée ancienne sous-tendant toute démarche scientifique à savoir celle du RASOIR D'OKHAM. Le principe du rasoir d'Okham consiste ici à dire que *toutes choses étant égales par ailleurs, une description plus simple sera toujours préférable à une description plus complexe*. Autrement dit, une description d'un système sera d'autant meilleure qu'elle reflètera plus simplement la richesse dudit système. Notre critère permet d'évaluer objectivement par une mesure dédiée les descriptions concurrentes d'un système flexionnel donné et de sélectionner l'une d'entre elles en vertu du critère objectif de l'ÉCONOMIE DE LA DESCRIPTION. Cette évaluation pourra alors s'appliquer pour toutes les langues que l'on voudrait décrire. Reste que l'interprétation des analyses et leur mise en situation ne relève ensuite plus de la morphologie formelle mais bien de la typologie. L'interprétation typologique pourra alors se faire sur des descriptions de systèmes flexionnels dont il aura été prouvé par la formalisation complète qu'elles sont cohérentes, par l'implémentation à grande échelle qu'elles sont complètes et ne négligent pas une partie des données et par l'évaluation quantitative qu'elles sont similairement satisfaisantes en termes d'économie descriptive.

## Variabilité et choix dans la conception d'analyses formelles de la flexion

Étudier la morphologie flexionnelle en adoptant une démarche à la fois formelle et typologique fait intervenir une TRIPLE VARIABILITÉ dans la conception des analyses : celle des SYSTÈMES ÉTUDIÉS, celle des CHOIX DU CADRE THÉORIQUE et celles des ANALYSES PARTICULIÈRES proposées pour les premiers à l'intérieur d'un modèle théorique donné.

Le premier type de variabilité, LA VARIABILITÉ DES SYSTÈMES FLEXIONNELS à travers les langues, appartient au domaine de la typologie linguistique. Elle relève d'une tâche empirique de collecte et de comparaison de données observables ainsi que de la mise en place d'un vocabulaire descriptif adapté à la phénoménologie rencontrée. Les primitives d'un modèle formel adapté à une démarche typologique doivent être posées à partir de vocabulaire descriptif. Dans cette thèse, nous avons choisi comme fondement conceptuel le paradigme de la TYPOLOGIE CANONIQUE élaboré par Greville G. Corbett et dont la

---

partie traitant de la flexion est notamment définie dans (Corbett, 2007a)<sup>7</sup>.

Le second paramètre dont dépend la conception d'un modèle formel est celui du CHOIX D'UN CADRE THÉORIQUE. Le cadre théorique choisi pour la conception d'un modèle formel détermine en effet la façon dont les données seront perçues et traitées. Dans la mesure où notre but est de rendre compte des données empiriques récoltées et catégorisées par le paradigme de la typologie canonique, il semble de première importance que le modèle formel ne contraigne pas l'analyse au point de ne plus pouvoir rendre compte des découvertes empiriques et des concepts typologiques préalables.

Nous avons choisi ainsi de développer notre modèle formel dans un CADRE RÉALISATIONNEL et non INCRÉMENTAL de la morphologie (Stump, 2001, chap. 1). La raison fondamentale et bien connue qui motive ce choix provient du problème de l'EXPONENCE MULTIPLE amplement discutée dans le paradigme de la typologie canonique (voir notamment (Baerman & Corbett, 2012)). Un exemple en est la forme *iront* en français en (0.1-a). Le trait du futur (FUT) y est exprimé dès le choix du radical supplétif *i-*. Il s'exprime ensuite de façon redondante par le suffixe temporel *-r-*, puis encore par l'exposant fusionnel *-ont*. Une approche incrémentale aurait des difficultés à rendre compte de cette exponence multiple dans la mesure où il ne serait pas possible de dire de *i-* ni de *-ont* qu'ils incrémentent la valeur du FUTUR dans *iront*, ce trait FUTUR étant déjà exprimé dès le choix du radical — et ce alors même qu'ils le font dans le cas de *fini-r-ont* à l'exemple (0.1-b).

- 0.1      a.    *i-r-ont*  
            aller.FUT-FUT-3PL.FUT
- b.    *fini-r-ont*  
            finir.-FUT-3PL.FUT

Dans la mesure où les approches réalisationnelles partent du principe que les traits réalisés par un mot-forme sont déjà pleinement spécifiés avant même l'application des règles de réalisation de la morphologie (qui se contentent de les exprimer par des exposants), elles ne contraignent en rien les cas d'exponence multiple et sont donc plus appropriées pour la conception d'un modèle formel pour la morphologie canonique qui pose l'exponence multiple au cœur des problématiques typologiques ayant trait à la flexion.

Notre modèle étant conçu en tant qu'outil pour la description de systèmes flexionnels à visée typologique, il semble par ailleurs essentiel que les descriptions proposées au sein de

---

7. Nous développons les arguments en faveur de ce choix au chapitre 3.

---

notre modèle soient évaluables quantitativement avant d'être confrontées aux descriptions des autres langues. Cette question rejoint le troisième type de variabilité auquel est soumis l'analyse formelle de la morphologie flexionnelle : celle des analyses à proprement parler.

Avant de pouvoir tirer des conclusions plus générales, notamment en termes de comparaison avec d'autres systèmes flexionnels, il convient alors de pouvoir en sélectionner une sur des critères objectifs. Un modèle formel doté d'un système d'implémentation efficace et facile d'emploi permet de comparer ces analyses par des mesures d'évaluation quantitative. Notre modèle s'accompagne ainsi d'UNE MÉTRIQUE OBJECTIVE DÉDIÉE À L'ÉVALUATION DE DESCRIPTIONS CONCURRENTES EN TERMES DE COMPACITÉ. Cette mesure s'appuie sur la notion de LONGUEUR DE DESCRIPTION MINIMALE (LDM) — *Minimum description Length (MDL)* (Rissanen, 1984) — empruntée à la théorie de l'information. Cette métrique constitue un moyen objectif externe pour contenir l'arbitraire inhérent aux descriptions constructives de sorte à les rendre utilisables pour des évaluations typologiques ultérieures.

## Le modèle PARSL

Dans cette thèse nous élaborons un modèle formel de la morphologie flexionnelle qui tente de répondre aux exigences définies ci-dessus. Notre modèle, nommé PARSL, présente les caractéristiques suivantes.

**Un modèle formel de la flexion canonique** Le concept de TYPOLOGIE CANONIQUE apparaît avec les travaux de Corbett (2003) et a pour but de permettre une meilleure compréhension de ce qui, dans le système formel que constitue une langue, différencie un idéal hypothétique, canonique, des différentes réalisations concrètes de phénomènes observables dans les langues. Une réalisation canonique se définit par une régularité maximale pour exprimer une ambiguïté minimale. En particulier, Corbett (2007a) donne une définition explicite de ce qui constitue la FLEXION CANONIQUE ainsi que les différents phénomènes irréguliers observables dans les paradigmes flexionnels des langues et appelés PHÉNOMÈNES NON CANONIQUES. Le modèle que nous proposons dans cette thèse a pour but de formaliser les critères définis par Corbett (2007a) afin de rendre compte de manière formelle de la phénoménologie établie dans le paradigme de la flexion canonique.

L'approche canonique de la morphologie permet d'évaluer qualitativement les déviations par rapport à un canon défini, grâce à des critères de canonicité précis. Mais quand

---

bien même le concept de canonicité est essentiellement une théorisation de la déviation, elle ne s'accompagne à ce jour d'aucun modèle formel permettant réellement de formaliser les déviations par rapport à ce canon. C'est entre autres ce manque que nous tentons de pallier en proposant dans cette thèse un nouveau modèle formel de la morphologie flexionnelle.

PARSLI a ainsi été spécifiquement conçu pour rendre compte des primitives qui sous-tendent l'approche canonique de la morphologie flexionnelle. Il propose une représentation explicite de la structure du paradigme de chaque entrée lexicale d'une langue donnée ainsi que de sa déviation par rapport au canon. La structure du paradigme y est conçue comme la conséquence des éléments constitutifs de la définition d'une entrée lexicale spécifique. L'étendue de chaque phénomène non canonique dans un paradigme donné peut ainsi être identifiée à partir des informations associées à l'entrée lexicale elle-même.

À partir de la formalisation que propose PARSLI, nous avons de surcroît défini des MESURES DE CANONICITÉ<sup>8</sup> permettant de rendre compte du degré de non-canonycité d'un paradigme ou d'une langue, et ce, explicitement, phénomène par phénomène. PARSLI permet ainsi d'étendre la puissance de l'approche canonique de la morphologie flexionnelle en lui offrant la possibilité non seulement de qualifier les différences d'un système flexionnel donné par rapport au canon, mais également de les quantifier.

**Un modèle dans la tradition MOT ET PARADIGME** Mais au-delà de la formalisation de la flexion canonique, PARSLI constitue également un modèle indépendant de morphologie flexionnelle, et en particulier du rapport entre LEXIQUE et STRUCTURES PARADIGMATIQUES, comme l'indique son nom acronymique PARSLI, signifiant « *PARadigm Shape and Lexicon Interface* ». Ce modèle est un modèle INFÉRENTIEL-RÉALISATIONNEL au sens de Stump (2001) s'inscrivant dans une approche de la morphologie de type MOT ET PARADIGME. Il se caractérise par une vision hautement structurée de la réalisation : une des innovations principales de PARSLI est ainsi la généralisation du concept de partitions (Pirrelli & Battista, 2000) ou d'espace thématique (Bonami & Boyé, 2003) au système exponentiel par l'introduction de la notion de ZONE FLEXIONNELLE.

**Une approche modulaire articulée de la grammaire** PARSLI repose sur une vision modulaire de la grammaire supposant à la fois une existence indépendante de la morphologie flexionnelle par rapport aux autres composantes de la grammaire, notamment de la syntaxe, et une interaction complexe avec ces dernières. En particulier, nous défendons

---

8. En plus de la MESURE DE COMPACTITÉ évoquée plus haut.

---

dans cette thèse une VISION MODULAIRE ARTICULÉE DE LA GRAMMAIRE. L'information linguistique est ainsi conçue comme étant répartie entre les différentes composantes de la grammaire (morphologie, syntaxe, phonologie, lexique...) et la modélisation optimale de cette répartition comme le résultat d'un critère externe objectif, à savoir celui de la compacité des descriptions indiqué plus haut.

En particulier,  $\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}$  comporte une théorisation de l'interaction de la composante flexionnelle avec un LEXIQUE STRUCTURÉ :  $\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}$  cherche ainsi à rendre compte explicitement du rapport entre la structuration des entrées lexicales dans le LEXIQUE FLEXIONNEL d'une part et l'organisation des opérations de la GRAMMAIRE FLEXIONNELLE. Il comporte notamment des définitions précises des notions de CLASSES FLEXIONNELLES, de l'ALLOMORPHIE RADICALE, de CATÉGORIE MORPHOLOGIQUE et d'ENTRÉE LEXICALE FLEXIONNELLE.

**Les outils d'implémentation et d'évaluation de descriptions** Le modèle formel  $\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}$  s'accompagne enfin d'un système d'implémentation Alexina $\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}$  spécialement adapté du formalisme lexical Alexina (Sagot, 2010). En particulier, nous avons pris soin de rendre le plus transparent possible le langage d'implémentation afin de permettre l'implémentation aisée des descriptions morphologiques développée en morphologie théorique. L'implémentation en Alexina $\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}$  des analyses conçues en  $\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}$  donne également accès aux outils de mesure de compacité mentionnés plus haut pour l'évaluation quantitative du degré d'économie d'une description donnée.

## Plan

Cette thèse est structurée en deux parties comportant respectivement trois et quatre chapitres.

La première partie constitue une introduction méthodologique à l'étude de la morphologie flexionnelle par des approches typologiques et formelles ainsi qu'un état de l'art général de ces approches dans ce domaine.

**Chapitre 1 : Méthodologie pour l'analyse formelle en morphologie : le cas des déponents latins** Dans le premier chapitre nous posons les choix méthodologiques sous-jacents au travail développé dans la suite de la thèse. À l'aide de l'exemple particulier des verbes déponents latins, nous illustrons la démarche générale de l'analyse morphologique allant de l'observation des données à l'analyse proprement dite en

---

passant par une représentation formelle explicite. Nous y étudions en particulier les deux propositions d'analyse des verbes déponents faites par Kiparsky (2005) d'une part et Hippisley (2007) de l'autre, en les confrontant aux exigences méthodologiques qu'il nous importe de définir. Nous y mettons notamment l'accent sur l'importance de la modélisation complète et de l'implémentation quasi-exhaustive comme garants de la cohérence globale d'une analyse. Nous illustrons la méthodologie générale par une proposition d'analyse originale des verbes déponents dont l'implémentation complète est indiquée disponible en ligne<sup>9</sup>. Ce chapitre présente également de façon informelle les caractéristiques générales du modèle formel que nous introduisons dans la seconde partie de cette thèse.

**Chapitre 2 : Approches formelles de la morphologie flexionnelle** Le deuxième chapitre présente les notions principales intervenant dans les CHOIX THÉORIQUES qui nous ont guidée dans la conception de notre modèle formel de la flexion. Nous présentons également les caractéristiques principales de deux MODÈLES INFÉRENTIELS-RÉALISATIONNELS dont  $\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}$  se rapproche le plus : *Paradigm Function Morphology* (PFM) (Stump, 2001, 2006) et *Network Morphology* (Corbett & Fraser, 1993; Brown & Hippisley, 2012) Enfin, à la section 2.2, nous mettons en place notre CONCEPTION MODULAIRE ARTICULÉE DE LA GRAMMAIRE qui jouera un rôle crucial dans les questions d'économie descriptive discutées au chapitre 6.

**Chapitre 3 : Qualifier la différence : la typologie canonique** Au chapitre 3, nous présentons les caractéristiques principales de trois familles d'approches typologiques de la flexion : les approches fondées sur les notions d'UNIVERSAUX et de PROTOTYPES issus de la tradition de Greenberg (1963), la MORPHOLOGIE NATURELLE (Mayerthaler, 1981; Wurzel, 1984; Dressler, 2000) et la TYPOLOGIE CANONIQUE (Corbett, 2003). Nous discutons les raisons qui nous ont amenée à choisir la typologie canonique comme fondement conceptuel du modèle formel que nous introduisons dans la seconde partie de cette thèse. Dans cette partie nous définissons également plus en détail les concepts sous-tendant l'approche canonique de la flexion (Corbett, 2007a) et présentons les différents phénomènes non canoniques qui ont été définis à partir des importantes études empiriques réalisées dans ce paradigme.

---

9. Sur la page personnelle de l'auteur.



---

La première partie pose ainsi l'ensemble des concepts employés par la typologie canonique pour rendre compte d'une façon qualitativement précise des types d'irrégularités paradigmatiques observables dans les langues du monde. Dans la seconde partie, nous introduisons notre proposition de formalisation de ces irrégularités au sein de  $\mathcal{PARSLI}$ , un modèle formel de la flexion dont les primitives reposent directement sur les concepts définis dans le paradigme de la flexion canonique.

**Chapitre 4 :  $\mathcal{PARSLI}$  : un modèle formel pour la morphologie flexionnelle** Ce chapitre est consacré à l'introduction du modèle formel de la morphologie flexionnelle  $\mathcal{PARSLI}$ . Il comporte, en un premier temps (sections 4.1 et 4.2), une présentation des propriétés saillantes et des primitives théoriques du modèle. Nous y décrivons notamment les notions de ZONE FLEXIONNELLE, de ZONE DE RADICAL et de ZONE THÉMATIQUE, ainsi que la REPRÉSENTATION MULTI-NIVEAUX de la flexion et les notions de SCHÈMES et SOUS-SCHÈMES flexionnels. Y figurent également la définition formelle des ENTRÉES LEXICALES FLEXIONNELLES, des CATÉGORIES MORPHOLOGIQUES, des différents phénomènes non canoniques comme l'HÉTÉROCLISE et les CLASSES FLEXIONNELLES ou le DÉCALAGE MORPHOSYNTAXIQUE représenté par la notion de FONCTION DE TRANSFERT.

À la section 4.4, nous présentons les MESURES DE CANONICITÉ que nous avons élaborées à partir de la formalisation en  $\mathcal{PARSLI}$  des définitions des phénomènes non canoniques issues de la typologie canonique. Nous montrons comment chacune d'entre elles permet d'indiquer quantitativement, phénomène par phénomène, le degré de non-canonlicité affiché par une entrée lexicale ou un système flexionnel donné.

**Chapitre 5 : Implémentations de descriptions** Dans ce chapitre, nous présentons le formalisme  $\text{Alexina}_{\mathcal{PARSLI}}$  que nous avons mis au point pour permettre l'implémentation de descriptions morphologiques  $\mathcal{PARSLI}$ . Après une discussion des avantages de l'implémentation, et notamment de l'implémentation à grande échelle (5.1), nous présentons à la section 5.2 les caractéristiques générales d' $\text{Alexina}_{\mathcal{PARSLI}}$ . À la section 5.3, nous illustrons à l'aide d'un exemple d'implémentation d'une description  $\mathcal{PARSLI}$  de la flexion du maltais les avantages concrets de l'implémentation à grande échelle en tant que garant de la cohérence et de la complétude d'une description morphologique.

**Chapitre 6 : Évaluations externes de descriptions morphologiques** Dans ce chapitre, nous nous penchons enfin sur la question de l'ÉVALUATION DE DESCRIPTIONS

---

FLEXIONNELLES CONCURRENTES par une métrique externe objective et les moyens de contenir la part de l'arbitraire inhérente aux approches constructives de la morphologie.

Nous décrivons l'élaboration de notre MÉTRIQUE DE COMPACTITÉ DE DESCRIPTION (section 6.1) et montrons son influence sur la question de la MAÎTRISE DE L'ARBITRAIRE DE LA SEGMENTATION et celle de la RÉPARTITION DE L'INFORMATION FLEXIONNELLE ENTRE LES DIFFÉRENTES COMPOSANTES DE LA GRAMMAIRE. En particulier, nous montrerons l'impact sur la compacité globale d'une description de la distribution de l'information entre la composante morphologique d'une part et les composantes du lexique (section 6.2), la (mor-)phonologie (section 6.3) et la syntaxe (section 6.4) de l'autre. Ce chapitre montrera ainsi également, à l'aide d'exemples du français, du maltais et du kurde sorani, le rôle de l'articulation entre les différentes composantes de la grammaire dans la conception d'une description globale visant à satisfaire le principe descriptif de l'optimisation de l'économie descriptive.

**Chapitre 7 : Mesures de canonicité et longueur de description** Ce dernier chapitre est dédié à l'application de l'ensemble de nos mesures à un exemple particulier. Nous y reprenons notre description de la flexion verbale du latin présentée au chapitre 1 et lui opposons une description alternative. Nous comparons ces deux descriptions à l'aide de notre métrique de compacité et leur appliquons également les mesures de canonicité mises en place au chapitre 4. Nous comparons enfin quantitativement ces deux descriptions à une implémentation d'une description  $\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}$  de l'analyse traditionnelle, en termes de compacité comme de canonicité.

## Lecture thématique

Le plan de cette thèse suit une progression conceptuelle depuis la conception d'un modèle jusqu'à son application en passant par son implémentation et ses métriques d'évaluation. Toutefois, une autre lecture, plus empirique, des sujets abordés dans ce travail est possible. Ainsi, les données empiriques principales abordées dans cette thèse sont les suivantes :

**Les verbes latins : les déponents dans le système verbal du latin** Une analyse des verbes déponents latins est proposée au chapitre 1. Nous y étudions les propriétés du passif en latin et montrons, à l'aide d'études empiriques sur le sujet, notamment (Flobert, 1967, 1975), que la passivation en latin gagne à être analysée comme une opération dérivationnelle

---

plutôt que flexionnelle. Nous en déduisons une analyse des verbes déponents comme des lexèmes hétéroclites se fléchissant à la fois selon la classe flexionnelle des verbes passifs canoniques et empruntant également certaines règles réalisationnelles issues de la classe des actifs canoniques.

Notre première analyse des verbes latins formulée au chapitre 1 repose sur une représentation à trois niveaux réalisationnels dans *PARSLI*. Au chapitre 7, nous la confrontons à une seconde analyse en quatre niveaux ainsi qu'à une formalisation implémentée de l'analyse traditionnelle et comparons ces trois analyses à la fois en termes de compacité et en termes de mesures de canonicité.

**Quatre descriptions concurrentes de la flexion verbale du français** Au chapitre 6, nous présentons et comparons en termes de compacité quatre descriptions concurrentes de la flexion verbale du français : celle de Bonami & Boyé (2003), celle du lexique *Lefff* décrite notamment dans (Sagot, 2010), une description générée automatiquement et une description nouvelle utilisant des concepts descriptifs nouveaux introduits par *PARSLI*.

**La flexion verbale maltaise** Nous présentons également deux analyses concurrentes de la flexion verbale du maltais. La première, au chapitre 5, présente une analyse essentiellement morphologique de la flexion des verbes du premier binyan qui s'appuie sur l'analyse de Camilleri (2011). Nous y confirmons par l'implémentation que la flexion des verbes maltais peut être représentée par un système morphologiquement complexe d'allomorphes radicales.

Au chapitre 6, nous proposons cependant une analyse alternative montrant notamment qu'une partie importante de la complexité morphologique observable dans l'allomorphie radicale des verbes maltais peut être réduite par un traitement morphologique s'appuyant sur un ensemble de onze règles et contraintes. Nous y comparons également les deux descriptions en termes de compacité globale.

**La flexion verbale du kurde sorani** Au chapitre 6, nous présentons enfin une extension de l'analyse des *MARQUES PERSONNELLES MOBILES* du kurde sorani de Samvelian (2007) et Bonami & Samvelian (2008). Nous montrons notamment que, contrairement à ce que pourraient en un premier temps laisser croire les données, leur comportement ne peut s'expliquer facilement par une analyse périphrastique des temps du parfait.

---

Les implémentations de toutes nos analyses flexionnelles sont disponibles en ligne <sup>10</sup>

## Collaborations scientifiques

Le travail présenté dans cette thèse s'appuie sur un certain nombre de collaborations scientifiques. Ci-dessous, nous détaillons comment ces collaborations s'articulent au contenu de la présente thèse.

Le développement du système d'implémentation Alexina<sub>PARSLI</sub> présenté au chapitre 5 est le fruit d'un travail commun avec Benoît Sagot (Alpage, INRIA & Université Paris Diderot). Le choix des primitives de l'implémentation découle de notre travail personnel, à savoir directement des primitives que nous avons posées dans notre modèle morphologique  $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}\mathcal{I}$ . Nous avons défini ensemble la façon de les représenter dans le formalisme Alexina<sub>PARSLI</sub>. L'implémentation et le développement des outils associés (outils de flexion, outils de calcul de compacité (cf. chapitre 6) ont été réalisés par Benoît Sagot.

Au chapitre 6, le travail présenté dans les sections 6.1 et 6.2 est également le fruit d'un travail commun avec Benoît Sagot. Il a notamment été publié dans (Sagot & Walther, 2011) et (Walther & Sagot, 2011) qui portent sur des expériences faites, à partir d'une redéfinition du formalisme Alexina (Sagot, 2010) à partir du modèle morphologique  $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}\mathcal{I}$  (Walther, 2011b), sur la comparaison de quatre descriptions du français à l'aide de mesures de longueur de descriptions minimales empruntées à la *Théorie de l'Information*. Le but de ce travail était de montrer la pertinence de  $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}\mathcal{I}$  pour concevoir des descriptions compactes de la morphologie flexionnelle d'une langue donnée. Ce travail commun s'appuie sur le développement de  $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}\mathcal{I}$  qui est de notre seul fait ainsi que sur l'adaptation du formalisme Alexina qui a été réalisée conjointement. En revanche, c'est Benoît Sagot qui a conçu seul la description originale du *Lefff* (Sagot, 2010) et adapté à l'écrit et implémenté la description de la flexion verbale proposée par Bonami & Boyé (2003) et la nôtre mettant en œuvre les notions développées dans  $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}\mathcal{I}$ . Benoît Sagot a également été responsable de la génération automatique de la description plate à partir du *Lefff* ainsi que de la conception et du développement des outils de mesure pour évaluer la compacité des descriptions concurrentes.

L'analyse implémentée de la flexion verbale du maltais présentée au chapitre 5 est le fruit d'une collaboration avec Maris Camilleri (Surrey Morphology Group (SMG), University

---

10. Sur la page personnelle de l'auteur.

---

of Surrey). Cette collaboration a été rendue possible grâce à une bourse du projet *NetWordS* de l'ESF (European Science Foundation) pour effectuer un séjour de recherche au SMG en janvier-février 2012.

La contribution respective des auteurs à ce travail est la suivante : l'analyse originelle des radicaux maltais a été conçue par Maris Camilleri et publiée dans (Camilleri, 2011). La formalisation de cette première description en  $\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}$  est de notre seul fait. Les nouvelles données traitées ont été extraites par nous-mêmes du *Maltese Language Resource Server* (Spagnol, 2011), filtrées par Maris Camilleri et encodées par nos soins. L'analyse de l'intégralité des données filtrées nous a conduit à concevoir conjointement une analyse des formes verbales complètes (section 5.3.3) dont la formalisation et l'implémentation en  $\text{Alexina}_{\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}}$  sont de notre seul fait. Cette analyse a été présentée conjointement dans (Camilleri & Walther, 2012).

Les données sur lesquelles portent notre étude de la flexion maltaise à la section 6.3 du chapitre 6 sont également issues de la ressource lexicale en ligne *Maltese Language Resource Server* (MLRS) publiée par Spagnol (2011) et de données complémentaires provenant de notre collaboration avec Maris Camilleri, *University of Surrey*. L'analyse (mor-)phonologique que nous présentons à la section 6.3.1 constitue une extension de l'analyse commune du chapitre 5, mais elle est de notre seul fait.

## Première partie

# État de l'art pour une approche typologique et formelle de la morphologie flexionnelle



# Méthodologie pour l'analyse formelle en morphologie : le cas des déponents latins

Dans ce premier chapitre, nous commençons par une étude de cas destinée à montrer le rôle de la formalisation en morphologie théorique et aussi l'importance de la méthodologie générale à suivre lors de l'analyse. L'exemple que nous étudions compare deux analyses récentes des verbes déponents latins pour finir par la proposition d'une nouvelle analyse des données latines qui suit les contraintes méthodologiques que nous nous fixons.

En latin, les verbes habituellement qualifiés de VERBES DÉPONENTS jouissent d'une notoriété ancienne et persistante depuis les études classiques ou descriptives (Bennett, 1907 ; Ernout & Thomas, 1953 ; Flobert, 1967, 1975) jusqu'aux approches théoriques ou formelles contemporaines (Börjars et al., 1996 ; Sadler & Spencer, 1998 ; Kiparsky, 2005 ; Hippisley, 2007 ; Stump, 2006 ; Baerman, 2007b). Les divergences pourtant affichées par les analyses existantes, tant dans leurs apports théoriques divers que dans leurs méthodologies variées, en font un excellent point d'accroche pour illustrer les contraintes méthodologiques en analyse morphologique qu'il nous importe de mettre en place dans ce chapitre introductif.

La célébrité de ces verbes tient très probablement à ce qu'ils affichent à première vue des propriétés contradictoires, relevées notamment par Flobert (1967) par la formule suivante :

« Les verbes déponents latins, passifs de forme mais fonctionnellement identiques aux actifs, constituent une énigme. »

(Flobert, 1967, p.568)



En effet, comme le montre l'exemple 1.1 de Cicéron (52 AEC)<sup>1</sup>, les déponents affichent les formes morphologiques de la voix passive tandis que leur emploi en syntaxe est celui de la voix active.

1.1      ..., ut eō                    tacente                    rēs                    ipsā  
                  que ce.ABL.SG taire.ABL.SG.PRS.PTCP.ACT chose.NOM.F.SG RFL.ABL.F.SG  
                  loqueretur.  
                  parler.3SG.IPFV.PST.SBJV.PASS  
                  « ...que même en taisant cela la chose parle d'elle-même. »

Notons, que le décalage apparent entre la forme morphologique passive et la fonction active dans la phrase de ce type de verbes a récemment servi de critère au sein des approches dites *canoniques* de la typologie pour forger le concept de *déponence* en tant que type d'*irrégularité* dans les systèmes flexionnels. En typologie canonique, le concept de déponence est explicitement défini par Baerman (2007b) comme un « *décalage entre forme et fonction* »<sup>2</sup>. Par ailleurs, ce concept de déponence y a ensuite été employé pour désigner autant le comportement des verbes déponents latins que l'emploi de formes du singulier pour exprimer le pluriel pour certains noms serbo-croates, et bien d'autres cas encore (Baerman, 2006). Dans la suite de ce chapitre, ainsi que dans les définitions des phénomènes flexionnels irréguliers qui feront l'objet des chapitres 3 et 4, nous mettrons en doute l'idée que les phénomènes étudiés par Baerman forment une classe cohérente. En attendant, nous continuons à employer le terme de déponence comme la caractéristique des seuls verbes déponents latins et nous parlerons du phénomène de DÉCALAGE (MORPHOSYNTAXIQUE) pour désigner le phénomène plus général théorisé par Baerman (2007b).

Nous définissons donc les VERBES DÉPONENTS comme des verbes qui emploient des formes morphologiques passives quand bien même ils n'expriment pas de valeur passive en tant que telle. Cette définition rejoint la définition traditionnelle des verbes déponents, à savoir des *verbes ayant* « déposé » *leur valeur passive*.

Il est à noter que les verbes déponents ne constituent pas un épiphénomène dans le système verbal global du latin. Xu et al. (2007) en relèvent un total de 543 dans le *Oxford Latin Dictionary* (Glare, 1983). Dans le lexique de 1 475 verbes de Prior & Wohlberg (1995)<sup>3</sup>

1. Citation complète : « *Nudavit se in sanctissimo templo, quoniam vita talis et civis et viri fidem non faciebat, ut eo tacente res ipsa loqueretur.* »

2. En anglais, « *a mismatch between form and function* ».

3. (Prior & Wohlberg, 1995) est un guide de conjugaison dont le but est de servir de référence pour déterminer la flexion d'un verbe latin donné. À ce titre, il ne contient que peu des verbes déverbaux dérivés

nous en avons relevé 144, c'est-à-dire 9,8% de l'inventaire complet des verbes.

Le but de ce chapitre est de comparer et d'évaluer en termes de méthodologie deux approches rendant compte des verbes déponents, à savoir l'approche lexicale de Kiparsky (2005) d'une part, et l'approche réalisationnelle de Hippiisley (2007) de l'autre. Ces deux approches diffèrent essentiellement sur deux plans.

**Différences théoriques :** Tandis que Hippiisley suit, apparemment par défaut, l'approche traditionnelle qui fait des formes passives un sous-paradigme du paradigme verbal complet, Kiparsky avance implicitement des données et des arguments en faveur d'une approche dérivationnelle de la passivation.

**Différences dans la modélisation :** La modélisation de Hippiisley est une modélisation complète en *Network Morphology* implémentée en DATR<sup>4</sup> pour un échantillon de verbes choisis, tandis que l'article de Kiparsky ne présente que les éléments principaux de la modélisation ; le lecteur n'y a pas accès à la formalisation dans son ensemble.

L'analyse que nous présenterons suite à la comparaison de ces deux analyses se donne les deux objectifs suivants :

**Enjeu théorique :** Compte tenu des données détaillées sur la valeur du passif morphologique en latin présentées notamment par Flobert (1967), nous commencerons par poser explicitement la question de la nature de l'opération morphologique que constitue en latin la passivation : flexion ou dérivation ? Nous verrons ensuite les conséquences de cette caractérisation sur l'analyse des verbes déponents.

**Enjeu formel :** Afin de pouvoir réellement rendre compte des données du latin, nous proposerons une formalisation qui se veut aussi explicite que celle de Hippiisley (2007). Nous la compléterons également par une implémentation avec lexique associé. Ainsi le lecteur peut-il lui-même constater le fonctionnement global de l'analyse en ayant accès à la formalisation complète en ligne<sup>5</sup>.

Notre analyse montre en particulier que la passivation, du moins en latin, gagne à être interprétée comme une opération dérivationnelle entre deux lexèmes appartenant à deux

---

dont le paradigme est identique à celui de leur base. Cela explique le nombre inférieur d'unités lexicales trouvées dans (Prior & Wohlberg, 1995) par rapport au *Oxford Latin Dictionary*, dictionnaire visant à répertorier l'ensemble des unités lexicales pertinentes, et ce autant pour les verbes en général que pour les verbes déponents en particulier.

4. Les fichiers de l'implémentation en question sont consultables en ligne sur le site du livre Brown & Hippiisley (2012) à la page <http://networkmorphology.as.uky.edu/theory/ch5latindtr> (valable au 11 juin 2013).

5. Sur la page personnelle de l'auteur, où le lecteur trouvera également le fichier contenant les 2 271 entrées lexicales traitées.

classes flexionnelles distinctes, faisant ainsi des verbes déponents des instances d'*hétéroclise* au sens de Stump (2006).

Ce chapitre est structuré comme suit. Nous commençons par une présentation sommaire de la flexion verbale latine. Nous présentons ensuite les propriétés essentielles du passif en latin, ainsi que les propriétés des paradigmes des verbes déponents, décrits notamment par Flobert (1967). Suit une confrontation des deux analyses mentionnées ci-dessus, Hippisley (2007) puis Kiparsky (2005). Enfin, nous terminons par la présentation de notre propre analyse dans laquelle nous mettrons l'accent sur la méthodologie adoptée.

## 1.1 Présentation générale du passif en latin

### 1.1.1 La flexion verbale du latin

#### 1.1.1.1 Présentation des paradigmes verbaux

En première analyse, les verbes latins présentent des paradigmes comportant jusqu'à 74 cases de formes actives et 69 cases de formes passives, soit un total de 143 cases possibles, dont une forme périphrastique pour l'actif et 31 pour le passif<sup>6</sup>.

Chaque forme verbale est construite par la combinaison d'un radical (cf. tableau 1.1 page 33) et d'un exposant (cf. tableaux 1.10 et 1.11 pages 45 à 46) et comporte par ailleurs une voyelle thématique (*a, e, i*), à moins d'appartenir à la classe athématique.

**Les radicaux** Chaque verbe (non défectif) affiche trois radicaux verbaux différents, représentés pour cinq verbes dans le tableau 1.1. Les radicaux sont chacun réservés à la formation d'une partie fixe du paradigme. La répartition des trois radicaux par rapport aux différents sous-paradigmes est ensuite indiquée par les mêmes teintes que celles du tableau 1.1 dans les tableaux 1.10 et 1.11 pages 45 à 46.

Traditionnellement, les trois radicaux verbaux latins sont qualifiés de radicaux *présent*, *passé* et *supin*. Cette appellation porte cependant à confusion en laissant croire à une répartition morphosyntaxiquement motivée des radicaux, alors même qu'il a été montré par Aronoff (1994) que le regroupement de formes partageant un même radical verbal en latin ne correspond pas à une telle motivation : elle représente une structure indépendante purement morphologique. Dans la suite de ce chapitre nous identifions les radicaux cités dans le tableau 1.1 par RAD1 pour le radical du présent, RAD2 pour le radical du passé et RAD3 pour le supin. Comme l'illustre le tableau 1.1, les radicaux RAD1, RAD2 et RAD3

---

6. L'adjectif verbal est commun à l'actif et au passif.

sont respectivement utilisés dans les formes de la première et deuxième personne du singulier de l'indicatif présent imperfectif et l'infinitif présent (pour le RAD1), la première personne du singulier du parfait (pour le RAD2) et le supin (pour le RAD3). Ces formes constituent les cinq formes de citation ou parties principales par lesquelles les verbes latins sont traditionnellement identifiés (voir par exemple *Le Grand Gaffiot* (Gaffiot & Flobert, 2000)).

CONJUGAISON	RAD1			RAD2	RAD3	TRADUCTION
	1SG.PRS.IPFV	2SG.PRS.IPFV	PRS.INF	1SG.PRS.PFV	SUP	
1 <sup>re</sup>	<i>am-ō</i>	<i>am-ā-s</i>	<i>am-ā-re</i>	<i>amāv-ī</i>	<i>amāt-um</i>	« aimer »
2 <sup>me</sup>	<i>del-e-ō</i>	<i>del-ē-s</i>	<i>del-ē-re</i>	<i>delēv-ī</i>	<i>delēt-um</i>	« détruire »
3 <sup>me</sup>	<i>leg-ō</i>	<i>leg-i-s</i>	<i>leg-e-re</i>	<i>leg-ī</i>	<i>lect-um</i>	« lire »
3 <sup>me</sup> mixte	<i>cap-i-ō</i>	<i>cap-i-s</i>	<i>cap-e-re</i>	<i>cep-ī</i>	<i>capt-um</i>	« saisir »
4 <sup>me</sup>	<i>aud-i-ō</i>	<i>aud-ī-s</i>	<i>aud-ī-re</i>	<i>audīv-ī</i>	<i>audīt-um</i>	« entendre »

TABLE 1.1 : Formes de citations avec indication des trois radicaux

**Classes flexionnelles** La flexion verbale latine s'organise par ailleurs traditionnellement (Ernout & Thomas, 1953 ; Sausy, 1977) autour de cinq classes flexionnelles, définies par des caractéristique morphologiques, à savoir par

- la nature de la voyelle thématique du verbe en question,
- la présence effective de cette voyelle thématique (qui peut tomber dans certaines cases).

Le tableau 1.1 présente un verbe de chacune de ces cinq classes verbales latines traditionnelles avec ses cinq parties principales traditionnelles.

Les voyelles thématiques sont également illustrées par les tableaux 1.2 à 1.5. Les tableaux montrent la présence (en rouge) d'une des trois voyelles thématiques, *a*, *e* et *i*) respectivement pour les verbes AMO « *aimer* », CONTINEO « *contenir* » et AUDIO « *entendre* », ainsi que l'absence de voyelle thématique pour le verbe athématique DUCO « *conduire, mener* ». Ces verbes correspondent respectivement à ce que les grammaires traditionnelles (Ernout & Thomas, 1953 ; Sausy, 1977) appellent les première, seconde, quatrième et troisième conjugaisons.

Par ailleurs, nous pouvons constater que les verbes AMO et CONTINEO se fléchissent selon un premier patron, qui comporte notamment la marque *-ba-* pour l'imparfait de l'indicatif (PST.IPFV) et un *-b-/-bi-/-bu-* pour le futur (FUT.IPFV), tandis que les verbes du type de DUCO et AUDIO suivent un autre patron avec *-eba-* à l'imparfait et *-e-* au futur.

Nous sommes donc, au-delà des différences thématiques, face à deux classes de désinences distinctes, partitionnant l'ensemble de verbes latins.

Aux quatre conjugaisons ainsi identifiées, s'ajoute par ailleurs une cinquième, habituellement qualifiée de *troisième mixte* comme le verbe ACCIPIO « *accepter* » illustré par le tableau 1.6. On remarque que ce type de verbes construit certaines de ces formes (cases colorées en vert foncé) comme les verbes à voyelle thématique *i*, tandis que les autres sont formées selon le patron des verbes athématiques<sup>7</sup> (cases colorées en vert clair). Nous les appellerons *verbes à thème mixte*. Les désinences de ce type de verbes suivent le second patron, à savoir celui de AUDIO et DUCO.

**Types de paradigmes** De façon orthogonale à leur répartition en cinq classes flexionnelles, les verbes latins diffèrent également par la structure globale de leur paradigme.

1) Les verbes transitifs comportent des formes actives autant que passives. Ils peuvent ainsi couvrir l'intégralité des 143 cases de paradigmes disponibles pour les verbes latins. Le lexique de Prior & Wohlberg (1995) comporte 1 112 verbes de ce type.

a) Parmi les paradigmes transitifs, les transitifs complets, cf. AMO « *aimer* », affichent l'inventaire complet des 143 formes verbales; Prior & Wohlberg (1995) listent 1 003 verbes transitifs non défectifs.

b) D'autres, défectifs au passif, comportent en général l'intégralité des 74 formes actives, mais sont partiellement défectifs pour le sous-paradigme passif. Le plus souvent ils comportent le seul sous-paradigme des formes de la troisième personne ( cf. CONSULO « *délibérer* »), parfois encore réduit aux seules formes du singulier (cf. VENIO « *venir* »), voire de l'imperfectif (cf. EGEO

---

7. Notons que les verbes athématiques comportent parfois une voyelle épenthétique s'insérant entre la finale consonantique de leur radical et une désinence à initiale consonantique. Un exemple de voyelle épenthétique est donné par les *-i-* dans les formes 2.SG.PRS.IPFV.IND *ducis* « *tu conduis* » et 3SG.PRS.IPFV.IND *ducit* « *il conduit* » du verbe DUCERE. On y remarque que la voyelle épenthétique est la même que la voyelle thématique observable pour ces mêmes cases pour le verbe AUDIO à thème *i* : *audīs* « *tu entends* » et *audit* « *il entend* ». Toutefois, le *i* de *audīs* est long, contrairement à celui de *ducis*, permettant dans ce cas de distinguer une voyelle épenthétique courte de la vraie voyelle thématique longue. Ce fait permet également d'identifier la forme *accipīs* « *tu acceptes* » de ACCIPIO comme relevant du thème athématique et comportant une voyelle épenthétique, et non du thème en *i* qui aurait comporté une voyelle longue : *\*accipīs*.

La voyelle thématique de la troisième personne du singulier de l'indicatif présent est en revanche toujours courte, comme le montrent notamment les formes *amat* « *il aime* » de AMARE et *continet* « *il contient* » de CONTINERE. Pour ces formes, une distinction entre voyelle thématique et voyelle épenthétique n'est donc pas aussi aisée. Dans les cases où il restait de l'ambiguïté quant à l'appartenance thématique d'un verbe mixte, nous avons fait le choix arbitraire de traiter ces cases comme des cases athématiques. Cela explique également l'attribution du thème athématique aux cases des verbes mixtes dont les formes sont construites à partir des radicaux RAD1 et RAD2 et qui n'affichent donc pas de distinction thématique.

AMO											
			ACT			PASS					
			IND	SG	SBJV	IND	SG	SBJV			
IPFV			IND	SG	PL	IND	SG	PL	IND	SG	PL
PFV	PRS	1	amā-ō	amā-mus	am-em	am-ēmus	am-ō-erim	amā-eris	amā-mur	am-er	am-ēmur
		2	amā-s	amā-tis	am-ēs	am-ētis	amā-eris	amā-eris (-re)	am-eris (-ēre)	am-eris (-ēre)	am-ēmī
		3	amā-t	amā-nt	am-et	am-ent	amā-erunt (-vère)	amā-ntur	am-entur	am-entur	am-ēmī
	PST	1	amā-bam	amā-bāmus	amā-rem	amā-rēmus	amā-erim	amā-bar	amā-bamur	amā-rem	amā-rēmur
		2	amā-bās	amā-bātis	amā-rēs	amā-rētis	amā-eris	amā-bāris (-bāre)	amā-bāris (-bāre)	amā-reris (-rēre)	amā-rēmī
		3	amā-bat	amā-bant	amā-ret	amā-rent	amā-erant	amā-bātur	amā-bātur	amā-rētur	amā-rēmī
	FUT	1	amā-bō	amā-bīmus	amā-rem	amā-rēmus	amā-erim	amā-bor	amā-bimur	amā-rem	amā-rēmur
		2	amā-bīs	amā-bītis	amā-ret	amā-rent	amā-eris	amā-beris (-bere)	amā-bimur	amā-reris (-rēre)	amā-rēmī
		3	amā-bit	amā-bunt	amā-ret	amā-rent	amā-erunt	amā-bitur	amā-buntur	amā-rētur	amā-rēmī
IMP	PRS	1	amā-ō	amā-mus	amā-erim	amā-erimus	amā-erim	amā-eris	amā-erimur	amā-eris	amā-erimur
		2	amā-s	amā-tis	amā-eris	amā-eritis	amā-eris	amā-eris	amā-eris	amā-eris	amā-eris
		3	amā-t	amā-nt	amā-erit	amā-erunt (-vère)	amā-erit	amā-erit	amā-erit	amā-erit	amā-erit
	PST	1	amā-bam	amā-bāmus	amā-erim	amā-erimus	amā-erim	amā-eris	amā-erimur	amā-eris	amā-erimur
		2	amā-bās	amā-bātis	amā-eris	amā-eritis	amā-eris	amā-eris	amā-eris	amā-eris	amā-eris
		3	amā-bat	amā-bant	amā-erit	amā-erunt	amā-erit	amā-erit	amā-erit	amā-erit	amā-erit
	FUT	1	amā-bō	amā-bīmus	amā-erim	amā-erimus	amā-erim	amā-eris	amā-erimur	amā-eris	amā-erimur
		2	amā-bīs	amā-bītis	amā-eris	amā-eritis	amā-eris	amā-eris	amā-eris	amā-eris	amā-eris
		3	amā-bit	amā-bunt	amā-erit	amā-erunt	amā-erit	amā-erit	amā-erit	amā-erit	amā-erit
INF	PRS	amā-re	amā-erim	amā-erim	amā-erimus	amā-erim	amā-eris	amā-erimur	amā-eris	amā-erimur	
	PST	amā-eris	amā-eritis	amā-eris	amā-eritis	amā-eris	amā-eris	amā-eris	amā-eris	amā-eris	
	FUT	amā-erit	amā-erunt	amā-erit	amā-erunt	amā-erit	amā-erit	amā-erit	amā-erit	amā-erit	
PTCP	PRS	amā-ens	amā-entes	amā-ens	amā-entes	amā-ens	amā-ens	amā-ens	amā-ens	amā-ens	
	PST	amā-eris	amā-eritis	amā-eris	amā-eritis	amā-eris	amā-eris	amā-eris	amā-eris	amā-eris	
	FUT	amā-erit	amā-erunt	amā-erit	amā-erunt	amā-erit	amā-erit	amā-erit	amā-erit	amā-erit	
GRV	PRS	amā-erim	amā-erimus	amā-erim	amā-erimus	amā-erim	amā-eris	amā-erimur	amā-eris	amā-erimur	
	PST	amā-eris	amā-eritis	amā-eris	amā-eritis	amā-eris	amā-eris	amā-eris	amā-eris	amā-eris	
	FUT	amā-erit	amā-erunt	amā-erit	amā-erunt	amā-erit	amā-erit	amā-erit	amā-erit	amā-erit	
GRD	PRS	amā-erim	amā-erimus	amā-erim	amā-erimus	amā-erim	amā-eris	amā-erimur	amā-eris	amā-erimur	
	PST	amā-eris	amā-eritis	amā-eris	amā-eritis	amā-eris	amā-eris	amā-eris	amā-eris	amā-eris	
	FUT	amā-erit	amā-erunt	amā-erit	amā-erunt	amā-erit	amā-erit	amā-erit	amā-erit	amā-erit	
SUP	PRS	amā-erim	amā-erimus	amā-erim	amā-erimus	amā-erim	amā-eris	amā-erimur	amā-eris	amā-erimur	
	PST	amā-eris	amā-eritis	amā-eris	amā-eritis	amā-eris	amā-eris	amā-eris	amā-eris	amā-eris	
	FUT	amā-erit	amā-erunt	amā-erit	amā-erunt	amā-erit	amā-erit	amā-erit	amā-erit	amā-erit	

TABLE 1.2 : Voyelle thématique *a* — exemple AMO « *aimer* »

CONTINEO

		ACT			PASS				
		SG	IND PL	ACT SG PL	IND SG PL	IND SG PL	SBJV SG PL		
IPFV	PRS	1	<i>contine-ō</i>	<i>contine-āmus</i>	<i>contine-am</i>	<i>contine-āmur</i>	<i>contine-ātur</i>	<i>contine-āmur</i>	
		2	<i>contine-s</i>	<i>contine-tis</i>	<i>contine-ās</i>	<i>contine-ritis (-re)</i>	<i>contine-ntur</i>	<i>contine-āntur</i>	
		3	<i>contine-t</i>	<i>contine-nt</i>	<i>contine-at</i>	<i>contine-tur</i>	<i>contine-ntur</i>	<i>contine-āntur</i>	
PST	1	<i>contine-bam</i>	<i>contine-bāmus</i>	<i>contine-rem</i>	<i>contine-bamur</i>	<i>contine-remur</i>	<i>contine-remur</i>		
		2	<i>contine-bas</i>	<i>contine-bātis</i>	<i>contine-rēs</i>	<i>contine-bāris (-bare)</i>	<i>contine-bāntur</i>	<i>contine-rētis (-rere)</i>	
		3	<i>contine-bat</i>	<i>contine-bant</i>	<i>contine-ret</i>	<i>contine-bātur</i>	<i>contine-bantur</i>	<i>contine-rētur</i>	
FUT	1	<i>contine-bo</i>	<i>contine-bimus</i>	<i>contine-rem</i>	<i>contine-bor</i>	<i>contine-bimur</i>	<i>contine-remur</i>		
		2	<i>contine-bis</i>	<i>contine-bitis</i>	<i>contine-beris (-bere)</i>	<i>contine-bentur</i>	<i>contine-bimur</i>	<i>contine-remur</i>	
		3	<i>contine-bit</i>	<i>contine-bunt</i>	<i>contine-bitur</i>	<i>contine-bitur</i>	<i>contine-buntur</i>	<i>contine-remur</i>	
PFV	PRS	1	<i>continu-ī</i>	<i>continu-imus</i>	<i>continu-erim</i>	<i>continu-erimus</i>	<i>conteni-ns sim</i>	<i>conteni-ns simus</i>	
		2	<i>continu-istī</i>	<i>continu-istis</i>	<i>continu-eris</i>	<i>continu-eritis</i>	<i>conteni-ns es</i>	<i>conteni-ns estis</i>	
		3	<i>continu-it</i>	<i>continu-erunt (-ere)</i>	<i>continu-erit</i>	<i>continu-erint</i>	<i>conteni-ns est</i>	<i>conteni-ns sunt</i>	
PST	1	<i>continu-eram</i>	<i>continu-erāmus</i>	<i>continu-essem</i>	<i>conteni-ns eram</i>	<i>conteni-ns eramus</i>	<i>conteni-ns essem</i>	<i>conteni-ns essemus</i>	
		2	<i>continu-erās</i>	<i>continu-erātis</i>	<i>continu-issēs</i>	<i>conteni-ns erās</i>	<i>conteni-ns erātis</i>	<i>conteni-ns essēs</i>	<i>conteni-ns essētis</i>
		3	<i>continu-erat</i>	<i>continu-erant</i>	<i>continu-isset</i>	<i>conteni-ns erat</i>	<i>conteni-ns erant</i>	<i>conteni-ns esset</i>	<i>conteni-ns essent</i>
FUT	1	<i>continu-erō</i>	<i>continu-erimus</i>	<i>continu-eris</i>	<i>conteni-ns erō</i>	<i>conteni-ns erimus</i>	<i>conteni-ns eris</i>	<i>conteni-ns eritis</i>	
		2	<i>continu-eris</i>	<i>continu-eritis</i>	<i>continu-erit</i>	<i>conteni-ns eris</i>	<i>conteni-ns eritis</i>	<i>conteni-ns eritis</i>	<i>conteni-ns eritis</i>
		3	<i>continu-erit</i>	<i>continu-erint</i>	<i>continu-erit</i>	<i>conteni-ns erit</i>	<i>conteni-ns erint</i>	<i>conteni-ns erit</i>	<i>conteni-ns erint</i>
IMP	PRS	2	<i>contine-te</i>	<i>contine-te</i>	<i>contine-te</i>	<i>contine-te</i>	<i>contine-te</i>	<i>contine-te</i>	
		2	<i>contine-to</i>	<i>contine-to</i>	<i>contine-to</i>	<i>contine-to</i>	<i>contine-to</i>	<i>contine-to</i>	
		3	<i>contine-to</i>	<i>contine-to</i>	<i>contine-to</i>	<i>contine-to</i>	<i>contine-to</i>	<i>contine-to</i>	
INF	PRS	<i>contine-re</i>			<i>contine-ri</i>				
		<i>continu-isse</i>			<i>conteni-ns/-a/-um esse</i>				
		<i>conteni-ns, -ntis</i>			<i>conteni-ns/-a/-um</i>				
PTCP	FUT	<i>conteni-ns, -ntis</i>			<i>conteni-ns/-a/-um</i>				
		<i>conteni-ns, -ntis</i>			<i>conteni-ns/-a/-um</i>				
		<i>conteni-ns, -ntis</i>			<i>conteni-ns/-a/-um</i>				
GRV GRD SUP	FUT	<i>conteni-ns, -ntis</i>			<i>conteni-ns/-a/-um</i>				
		<i>conteni-ns, -ntis</i>			<i>conteni-ns/-a/-um</i>				
		<i>conteni-ns, -ntis</i>			<i>conteni-ns/-a/-um</i>				

TABLE 1.3 : Voyelle thématique e — exemple CONTINEO « conteni-ri »

		AUDIO								
		ACT			PASS					
		IND		SBJV		IND		SBJV		
		SG	PL	SG	PL	SG	PL	SG	PL	
IPFV	PRS	1	<i>audi-mus</i>	<i>audi-am</i>	<i>audi-āmus</i>	<i>audi-or</i>	<i>audi-mur</i>	<i>audi-ar</i>	<i>audi-āmur</i>	
		2	<i>audi-tis</i>	<i>audi-ās</i>	<i>audi-ātis</i>	<i>audi-ris (-re)</i>	<i>audi-mini</i>	<i>audi-āris (-āre)</i>	<i>audi-āmini</i>	
		3	<i>audi-unt</i>	<i>audi-at</i>	<i>audi-ant</i>	<i>audi-tur</i>	<i>audi-untur</i>	<i>audi-ātur</i>	<i>audi-āntur</i>	
	PST	1	<i>audi-ēbam</i>	<i>audi-ēbamus</i>	<i>audi-rem</i>	<i>audi-remus</i>	<i>audi-ēbar</i>	<i>audi-ēbāmur</i>	<i>audi-rer</i>	<i>audi-rēmur</i>
		2	<i>audi-ēbās</i>	<i>audi-ēbātis</i>	<i>audi-rēs</i>	<i>audi-rētis</i>	<i>audi-ēbāris (-ēbāre)</i>	<i>audi-ēbāmini</i>	<i>audi-rēris (-rēre)</i>	<i>audi-rēmini</i>
		3	<i>audi-ēbant</i>	<i>audi-ēbant</i>	<i>audi-ret</i>	<i>audi-rent</i>	<i>audi-ēbātur</i>	<i>audi-ēbantur</i>	<i>audi-rētur</i>	<i>audi-rēntur</i>
	FUT	1	<i>audi-am</i>	<i>audi-ēmus</i>			<i>audi-ar</i>	<i>audi-ēmur</i>		
		2	<i>audi-ēs</i>	<i>audi-ētis</i>			<i>audi-eris (-ēre)</i>	<i>audi-ēmini</i>		
		3	<i>audi-et</i>	<i>audi-ent</i>			<i>audi-ētur</i>	<i>audi-entur</i>		
PFV	PRS	1	<i>audio-i</i>	<i>audio-erim</i>	<i>audio-erimus</i>	<i>audio-us sum</i>	<i>audio-i sumus</i>	<i>audio-us sim</i>	<i>audio-i simus</i>	
		2	<i>audio-isti</i>	<i>audio-eris</i>	<i>audio-eritis</i>	<i>audio-us es</i>	<i>audio-i estis</i>	<i>audio-us sis</i>	<i>audio-i sitis</i>	
		3	<i>audio-it</i>	<i>audio-erunt (-ēre)</i>	<i>audio-erit</i>	<i>audio-us est</i>	<i>audio-i sunt</i>	<i>audio-us sit</i>	<i>audio-i sint</i>	
	PST	1	<i>audio-eram</i>	<i>audio-erāmus</i>	<i>audio-issem</i>	<i>audio-issemus</i>	<i>audio-us eram</i>	<i>audio-i erāmus</i>	<i>audio-us essem</i>	<i>audio-i essemus</i>
		2	<i>audio-erās</i>	<i>audio-erātis</i>	<i>audio-isses</i>	<i>audio-issetis</i>	<i>audio-us erās</i>	<i>audio-i erātis</i>	<i>audio-us esēs</i>	<i>audio-i esētis</i>
		3	<i>audio-erat</i>	<i>audio-erant</i>	<i>audio-isset</i>	<i>audio-issent</i>	<i>audio-us erat</i>	<i>audio-i erant</i>	<i>audio-us esset</i>	<i>audio-i essent</i>
	FUT	1	<i>audio-erō</i>	<i>audio-erimus</i>			<i>audio-us erō</i>	<i>audio-i erimus</i>		
		2	<i>audio-eris</i>	<i>audio-eritis</i>			<i>audio-us eris</i>	<i>audio-i eritis</i>		
		3	<i>audio-erit</i>	<i>audio-erint</i>			<i>audio-us erit</i>	<i>audio-i erint</i>		
IMP	PRS	2	<i>audi-te</i>			<i>audi-re</i>	<i>audi-mini</i>			
	FUT	2	<i>audi-to</i>			<i>audi-tor</i>	<i>audi-ntor</i>			
	3	<i>audi-to</i>				<i>audi-tor</i>	<i>audi-ntor</i>			
INF	PRS		<i>audi-re</i>					<i>audi-rī</i>		
	PST		<i>audio-isse</i>					<i>audi-rī</i>		
	FUT		<i>audio-ūrus/-ūra/-ūrum esse</i>					<i>audi-us/-a/-um esse</i>		
PTCP	PRS		<i>audi-ēns, -ēntis</i>							
	PST		<i>audi-ūrus/-ūra/-ūrum</i>							
	FUT		<i>audi-ūrus/-ūra/-ūrum</i>							
GRV	PRS		<i>audi-ūrus/-ūra/-ūrum</i>							
	PST		<i>audi-endū/-endū/-endum</i>							
	FUT		<i>audi-endus/-enda/-endum</i>							
GRD	PRS		<i>audi-ūrus/-ūra/-ūrum</i>							
	FUT		<i>audi-ūrus/-ūra/-ūrum</i>							
SUP	PRS		<i>audi-ūrus/-ūra/-ūrum</i>							
	FUT		<i>audi-ūrus/-ūra/-ūrum</i>							

TABLE 1.4 : Voyelle thématique *i* — exemple AUDIO « entendre »



		DUCCO				
		ACT		PASS		
		SG	PL	SG	PL	
IPFV	PRS	1	<i>duc-ō</i>	<i>duc-am</i>	<i>duc-or</i>	<i>duc-antur</i>
		2	<i>duc-s</i>	<i>duc-as</i>	<i>duc-eris (-re)</i>	<i>duc-amini</i>
		3	<i>duc-t</i>	<i>duc-at</i>	<i>duc-tur</i>	<i>duc-untur</i>
PST	PRS	1	<i>duc-ebam</i>	<i>duc-erem</i>	<i>duc-ebar</i>	<i>duc-ebamur</i>
		2	<i>duc-ebās</i>	<i>duc-erēs</i>	<i>duc-ebāris (-ebāre)</i>	<i>duc-ebamini</i>
		3	<i>duc-ebat</i>	<i>duc-eret</i>	<i>duc-ebātur</i>	<i>duc-ebantur</i>
FUT	FUT	1	<i>duc-am</i>	<i>duc-erim</i>	<i>duc-ar</i>	<i>duc-eris (-ere)</i>
		2	<i>duc-ēs</i>	<i>duc-eritis</i>	<i>duc-eris (-ere)</i>	<i>duc-erimini</i>
		3	<i>duc-et</i>	<i>duc-erint</i>	<i>duc-entur</i>	<i>duc-entur</i>
PFV	PRS	1	<i>dux-i</i>	<i>dux-erim</i>	<i>duct-us sum</i>	<i>duct-i sumus</i>
		2	<i>dux-isti</i>	<i>dux-eris</i>	<i>duct-us es</i>	<i>duct-i estis</i>
		3	<i>dux-it</i>	<i>dux-erit</i>	<i>duct-us est</i>	<i>duct-i sunt</i>
PST	PST	1	<i>dux-eram</i>	<i>dux-essem</i>	<i>duct-us eram</i>	<i>duct-i erāmus</i>
		2	<i>dux-erās</i>	<i>dux-essēs</i>	<i>duct-us erās</i>	<i>duct-i erātis</i>
		3	<i>dux-erat</i>	<i>dux-isset</i>	<i>duct-us erat</i>	<i>duct-i erant</i>
FUT	FUT	1	<i>dux-erō</i>	<i>dux-erimus</i>	<i>duct-us erō</i>	<i>duct-i erimus</i>
		2	<i>dux-eris</i>	<i>dux-eritis</i>	<i>duct-us eris</i>	<i>duct-i eritis</i>
		3	<i>dux-erit</i>	<i>dux-erint</i>	<i>duct-us erit</i>	<i>duct-i erunt</i>
IMP	PRS	2	<i>duc</i>	<i>duc-te</i>	<i>duce-re</i>	<i>duct-mini</i>
		2	<i>duc-to</i>	<i>duc-tōte</i>	<i>duc-tor</i>	<i>duc-tor</i>
		3	<i>duc-to</i>	<i>duc-tōte</i>	<i>duc-tor</i>	<i>duc-tor</i>
INF	PRS	<i>duc-re</i>		<i>duc-i</i>		
		<i>dux-isse</i>		<i>duct-us/-a/-um esse</i>		
		<i>duct-urus/-ura/-urum esse</i>		<i>duct-us/-a/-um esse</i>		
PTCP	PRS	<i>duc-ens, -entis</i>		<i>duct-us/-a/-um</i>		
		<i>duct-urus/-ura/-urum</i>		<i>duct-us/-a/-um</i>		
		<i>duc-endi/-endō/-endum</i>		<i>duc-endus/-enda/-endum</i>		
GRV	PST	<i>duc-urus/-ura/-urum</i>		<i>duc-urus/-ura/-urum</i>		
		<i>duc-endi/-endō/-endum</i>		<i>duc-endus/-enda/-endum</i>		
		<i>duc-urus/-ura/-urum esse</i>		<i>duc-urus/-ura/-urum esse</i>		
GRD	PST	<i>duc-urus/-ura/-urum</i>		<i>duc-urus/-ura/-urum</i>		
		<i>duc-endi/-endō/-endum</i>		<i>duc-endus/-enda/-endum</i>		
		<i>duc-urus/-ura/-urum esse</i>		<i>duc-urus/-ura/-urum esse</i>		
SUP	PST	<i>duc-urus/-ura/-urum</i>		<i>duc-urus/-ura/-urum</i>		
		<i>duc-endi/-endō/-endum</i>		<i>duc-endus/-enda/-endum</i>		
		<i>duc-urus/-ura/-urum esse</i>		<i>duc-urus/-ura/-urum esse</i>		

TABLE 1.5 : Verbes athématiques — exemple DUCCO « *conduire, mener* »

		ACCIPIO										
		ACT			PASS							
IPFV	PRS PST FUT	SG IND PL	SG SBJV PL	SG IND PL	SG IND PL	SG SBJV PL						
							1	2	3	1	2	3
PFV	PRS	1	accip-ō	accip-ō	accip-ō	accip-ō	accip-ō	accip-ō	accip-ō	accip-ō	accip-ō	accip-ō
		2	accip-s	accip-s	accip-s	accip-s	accip-s	accip-s	accip-s	accip-s	accip-s	accip-s
		3	accip-t	accip-t	accip-t	accip-t	accip-t	accip-t	accip-t	accip-t	accip-t	accip-t
	PST	1	accip-ēbam	accip-ēbam	accip-ēbam	accip-ēbam	accip-ēbam	accip-ēbam	accip-ēbam	accip-ēbam	accip-ēbam	accip-ēbam
		2	accip-ēbas	accip-ēbas	accip-ēbas	accip-ēbas	accip-ēbas	accip-ēbas	accip-ēbas	accip-ēbas	accip-ēbas	accip-ēbas
		3	accip-ēbat	accip-ēbat	accip-ēbat	accip-ēbat	accip-ēbat	accip-ēbat	accip-ēbat	accip-ēbat	accip-ēbat	accip-ēbat
	FUT	1	accip-am	accip-am	accip-am	accip-am	accip-am	accip-am	accip-am	accip-am	accip-am	accip-am
		2	accip-ēs	accip-ēs	accip-ēs	accip-ēs	accip-ēs	accip-ēs	accip-ēs	accip-ēs	accip-ēs	accip-ēs
		3	accip-et	accip-et	accip-et	accip-et	accip-et	accip-et	accip-et	accip-et	accip-et	accip-et
IMP	PRS	1	accip-ī	accip-ī	accip-ī	accip-ī	accip-ī	accip-ī	accip-ī	accip-ī	accip-ī	accip-ī
		2	accip-istī	accip-istī	accip-istī	accip-istī	accip-istī	accip-istī	accip-istī	accip-istī	accip-istī	accip-istī
		3	accip-it	accip-it	accip-it	accip-it	accip-it	accip-it	accip-it	accip-it	accip-it	accip-it
	PST	1	accip-eram	accip-eram	accip-eram	accip-eram	accip-eram	accip-eram	accip-eram	accip-eram	accip-eram	accip-eram
		2	accip-erās	accip-erās	accip-erās	accip-erās	accip-erās	accip-erās	accip-erās	accip-erās	accip-erās	accip-erās
		3	accip-erat	accip-erat	accip-erat	accip-erat	accip-erat	accip-erat	accip-erat	accip-erat	accip-erat	accip-erat
	FUT	1	accip-erō	accip-erō	accip-erō	accip-erō	accip-erō	accip-erō	accip-erō	accip-erō	accip-erō	accip-erō
		2	accip-eris	accip-eris	accip-eris	accip-eris	accip-eris	accip-eris	accip-eris	accip-eris	accip-eris	accip-eris
		3	accip-erit	accip-erit	accip-erit	accip-erit	accip-erit	accip-erit	accip-erit	accip-erit	accip-erit	accip-erit
INF	2	accipe	accipe	accipe	accipe	accipe	accipe	accipe	accipe	accipe	accipe	
	2	accip-to	accip-to	accip-to	accip-to	accip-to	accip-to	accip-to	accip-to	accip-to	accip-to	
	3	accip-to	accip-to	accip-to	accip-to	accip-to	accip-to	accip-to	accip-to	accip-to	accip-to	
PTCP	PRS	accipe-re	accipe-re	accipe-re	accipe-re	accipe-re	accipe-re	accipe-re	accipe-re	accipe-re	accipe-re	
	PST	accip-isse	accip-isse	accip-isse	accip-isse	accip-isse	accip-isse	accip-isse	accip-isse	accip-isse	accip-isse	
	FUT	accip-urus/-ūra/-ūrum esse	accip-urus/-ūra/-ūrum esse	accip-urus/-ūra/-ūrum esse	accip-urus/-ūra/-ūrum esse	accip-urus/-ūra/-ūrum esse	accip-urus/-ūra/-ūrum esse	accip-urus/-ūra/-ūrum esse	accip-urus/-ūra/-ūrum esse	accip-urus/-ūra/-ūrum esse	accip-urus/-ūra/-ūrum esse	
GRV	PRS	accip-ens, -ēntis	accip-ens, -ēntis	accip-ens, -ēntis	accip-ens, -ēntis	accip-ens, -ēntis	accip-ens, -ēntis	accip-ens, -ēntis	accip-ens, -ēntis	accip-ens, -ēntis	accip-ens, -ēntis	
	PST	accip-urus/-ūra/-ūrum	accip-urus/-ūra/-ūrum	accip-urus/-ūra/-ūrum	accip-urus/-ūra/-ūrum	accip-urus/-ūra/-ūrum	accip-urus/-ūra/-ūrum	accip-urus/-ūra/-ūrum	accip-urus/-ūra/-ūrum	accip-urus/-ūra/-ūrum	accip-urus/-ūra/-ūrum	
	FUT	accip-endus/-enda/-endum	accip-endus/-enda/-endum	accip-endus/-enda/-endum	accip-endus/-enda/-endum	accip-endus/-enda/-endum	accip-endus/-enda/-endum	accip-endus/-enda/-endum	accip-endus/-enda/-endum	accip-endus/-enda/-endum	accip-endus/-enda/-endum	
SUP	1	accip-um/-ū	accip-um/-ū	accip-um/-ū	accip-um/-ū	accip-um/-ū	accip-um/-ū	accip-um/-ū	accip-um/-ū	accip-um/-ū	accip-um/-ū	
	2	accip-ior	accip-ior	accip-ior	accip-ior	accip-ior	accip-ior	accip-ior	accip-ior	accip-ior	accip-ior	
	3	accip-issim	accip-issim	accip-issim	accip-issim	accip-issim	accip-issim	accip-issim	accip-issim	accip-issim	accip-issim	

TABLE 1.6 : Verbes à thème mixte — exemple ACCIPIO « accepter »

« *manquer* »). Dans le lexique de Prior & Wohlberg (1995), on trouve 109 entrées verbales transitives défectives.

- 2) Les verbes intransitifs réguliers n'affichent que des formes morphologiques actives, à savoir 74 formes, cf. FAVEO « *préférer* ». Ils sont au nombre de 215 dans le lexique de Prior & Wohlberg (1995), dont 193 avec des paradigmes complets et 22 défectifs.
- 3) Les verbes déponents, de même que quelques rares verbes dits *passiva tantum* (Flobert, 1967), affichent uniquement les formes finies passives. Ces verbes ont la particularité de posséder, en plus des 69 formes passives, quelques rares formes actives, à savoir les participes présent et futur, les gérondifs et le supin. Ces verbes sont au nombre de 141 chez Prior & Wohlberg (1995) dont 137 avec des paradigmes complets et 4 défectifs.
- 4) Enfin, il existe aussi deux types de verbes assez rares généralement qualifiés de *semi-déponents*. Nous les désignerons par la suite de verbes déponents mixtes. Le premier type de verbes déponents mixtes (désormais DM1) présente des formes actives à l'imperfectif et passives au perfectif; il en existe en tout six : AUDEO « *oser* », GAUDEO « *se réjouir* », SOLEO « *avoir coutume* », FIDO « *avoir confiance* », CONFIDO « *avoir confiance* », DIFFIDO « *n'avoir pas confiance* » (Sausy, 1977). Le deuxième type (désormais DM2) présente des formes passives à l'imperfectif et actives au perfectif; il est bien moins fréquent encore au sens qu'on en trouve qu'un seul cas<sup>8</sup> REVERTOR « *revenir* » (Give, 1992).
- 5) Enfin, le latin possède par ailleurs toute une série de verbes défectifs variés, cf. LICET « *il est permis* », AFFOR « *adresser* », IMPENDEO « *surplomber* ». Ces verbes, au nombre de 135 chez Prior & Wohlberg (1995), ont des paradigmes extrêmement irréguliers, certains étant des verbes déponents défectifs (Flobert, 1967; Prior & Wohlberg, 1995). De ces derniers Prior & Wohlberg (1995) en comptent 4.

Le tableau 1.7 donne une indication simplifiée sur les proportions des types de paradigmes verbaux que l'on peut rencontrer en latin<sup>9</sup>.

Outre les problèmes particuliers liés à la défectivité répandue (9,2% des entrées dans (Prior & Wohlberg, 1995)) dans les paradigmes latins, les structures des paradigmes latins affichent donc une diversité notable. Dans ce chapitre, notre étude des verbes déponents

---

8. Ce cas est par ailleurs problématique dans la mesure où il existe également un verbe REVERTO synonyme de REVERTOR qui se comporte exactement comme les verbes intransitifs réguliers, cf. FAVEO « *préférer* » ci-dessus.

9. Les comptes ont été fait à partir des 1 474 verbes du lexique de Prior & Wohlberg (1995) auquel nous avons ajouté le déponent mixte REVERTOR « *revenir* ».

TYPE	COMPLETS	DÉFECTIFS	TOTAL
TOTAL	1 340 (90,8%)	135 (9,2%)	1 475 (100,0%)
TRANSITIFS	1 003 (68,0%)	109 (7,4%)	1 112 (75,4%)
INTRANSITIFS	193 (13,1%)	22 (1,2%)	215 (14,6%)
VERBES DÉPONENTS	137 (9,3%)	4 (0,3%)	141 (9,6%)
VERBES DÉPONENTS MIXTES	7 (0,5%)	– (0%)	7 (0,5%)

TABLE 1.7 : Répartition des types de verbes en latin (données de Prior & Wohlberg (1995))

s'intéresse en particulier au rapport entre la structure des paradigmes, le marquage morphologique de la voix et l'alignement des formes actives/passives avec les valeurs, notamment syntaxiques, de la voix passive.

**Type de marquage** Le latin dispose d'un marquage morphologique explicite, essentiellement affixal, de la voix. Au passif, les désinences personnelles habituelles sont remplacées par des désinences spécifiques. Un extrait de l'alternance désinentielle ainsi obtenue est illustrée dans le tableau 1.8 pour le présent de l'indicatif (imperfectif). Notons aussi que cette alternance n'est valable que pour les formes imperfectives, les formes perfectives passives étant périphrastiques (participe passé plus auxiliaire SUM « être »).

Un aperçu global du marquage de l'actif et du passif est par ailleurs donné dans les tableaux récapitulatifs 1.10 et 1.11 pages 45 à 46.

	ACTIF	PASSIF
1SG	–ō	–or
2.SG	–s	–ris ou –re
3SG	–t	–tur
1PL	–mus	–mur
2PL	–tis	–mini
3PL	–nt	–ntur

TABLE 1.8 : Désinences actives et passives au présent de l'indicatif imperfectif

Afin de distinguer le marquage morphologique et la valeur syntaxique des termes *actif* et *passif*, nous utilisons désormais le terme *m-actif* (respectivement *m-passif*) pour les formes construites à partir des marques morphologiques actives (respectivement passives). Nous appelons *s-passif* (respectivement *s-actif*) la notion syntaxique de passif (respectivement d'actif) définie plus bas.

**Alignement morphosyntaxique** Compte tenu des données présentées jusqu'ici nous observons qu'il y a quatre types d'alignement morphosyntaxiques possibles en latin. Ces alignements sont résumés de façon simplifiée dans le tableau 1.9. Il n'est notamment pas précisé que les verbes déponents et les passiva tantum possèdent également quelques formes m-actives, à savoir les formes pour les participes actifs, le gérondif et le supin.

Notons que la différence principale entre les verbes déponents et les passiva tantum réside dans le fait que les premiers présentent un alignement morphosyntaxique m-passif/s-actif lorsque les seconds alignent le m-passif avec le s-passif. Autrement dit, la différence entre ces deux types de verbes n'est donc pas morphologique (leurs paradigmes sont identiques) mais bien syntaxique : les verbes déponents apparaissent dans une construction active là où les passiva tantum apparaissent dans une construction passive.

Les rares déponents mixtes combinent exceptionnellement deux alignements distincts dans deux parties distinctes de leur paradigme<sup>10</sup>.

TYPE VERBAL	S-ACTIF	S-PASSIF
TRANSITIFS	m-actif	m-passif
INTRANSITIFS	m-actif	
VERBES DÉPONENTS	m-passif	
PASSIVA TANTUM		m-passif

TABLE 1.9 : Alignements possibles entre la morphologie verbale et la syntaxe pour les verbes latins

### 1.1.2 Le passif en latin

Les caractéristiques formelles des paradigmes verbaux latins ayant été établies, il convient désormais de décrire plus en détail les valeurs associées à l'actif et au passif en latin. Nous nous appuyons pour cela sur l'étude détaillée de Flobert (1967). La question qui nous intéresse ici particulièrement est celle de la nature morphologique de la passivation en latin. En étudiant les caractéristiques du passif latin, nous nous concentrerons ainsi tout particulièrement sur la question de savoir s'il résulte d'une opération flexionnelle ou dérivationnelle. Nous verrons que quand bien même l'interprétation flexionnelle du passif constitue l'approche la plus commune, l'analyse dérivationnelle semble en fait bien plus appropriée compte tenu des propriétés morphologiques et sémantiques du passif latin.

10. Compte tenu du caractère très ponctuel de ce phénomène, le 0,5% de verbes déponents mixtes n'est pas repris en compte dans ce tableau récapitulatif.

D'après Siewierska (2012), une construction peut être classée parmi les constructions passives si elle affiche les cinq propriétés suivantes :

- elle contraste avec une autre construction, dite construction active ;
- le sujet de l'actif correspond à un syntagme oblique optionnel au passif ou n'est pas explicitement exprimé ;
- le sujet du passif, s'il existe, correspond à l'objet direct de l'actif ;
- la construction est pragmatiquement contrainte par rapport à l'actif ;
- la construction affiche un marquage morphologique particulier sur le verbe.

Elle cite l'exemple 1.2 de (Ashton, 1947, p. 224) pour le swahili comme exemple prototypique de construction passive. Le sujet *Hamisi* de la phrase active (1.2-a) correspond à l'oblique facultatif *na Hamisi* dans la phrase (1.2-b), tandis que l'objet direct *chakula* de la phrase (1.2-a) correspond au sujet dans la phrase 1.2. Morphologiquement, le passif est marqué sur le verbe par la présence de *-w-*.

- 1.2 Swahili (Ashton, 1947, p. 224)
- a. *Hamisi a-li-pik-a chakula*  
 Hamisi 3SG-PST-cuisiner-IND nourriture  
 « *Hamisi a cuisiné (de) la nourriture.* »
- b. *chakula ki-li-pik-w-a (na Hamisi)*  
 nourriture 3SG-PST-cuisiner-PASS-IND par Hamisi  
 « *La nourriture a été cuisinée (par Hamisi).* »

En latin, les formes utilisées sont celles du tableau 1.11 ; l'argument syntaxique optionnel dénotant le sujet de la construction active correspondante est, dans les cas où il est présent, marqué par la préposition *a(b) + ABL*, voire par le seul ablatif (Flobert, 1967, 1975).

Notons en particulier que l'existence d'un argument syntaxique optionnel marqué par la préposition *a(b) + ABL* n'est pas selon Flobert (1967, 1975) une propriété exclusive du passif, certains verbes actifs comme *CADO* « *tomber* », *PEREO* « *périr* », *VAPULO* « *recevoir des coups* » et *VENIO* « *venir* » et certains verbes déponents comme *MOROR* « *mourir* », *NASCOR* « *naître* » et *ORIOR* « *prier* » pouvant également apparaître dans des constructions de ce type. Il n'est pas pertinent de parler de construction passive dans la mesure où la construction avec *a(b) + ABL* n'alterne pas avec une construction à accusatif (Flobert, 1975, p. 29).

### 1.1.2.1 Le marquage du s-passif en latin

Pour l'imperfectif, le marquage du passif se fait par le remplacement des désinences m-actives (cf. tableau 1.10) par des désinences m-passives (cf. tableau 1.11). On remarque également que les formes perfectives dans le tableau 1.11 sont périphrastiques. La valeur de la construction obtenue est alors en général celle de l'expression du s-passif, comme l'illustre par l'exemple 1.3 chez (Flobert, 1967, p. 571).

1.3 Latin (Flobert, 1967, p. 571)

- a. *Mercurius Sosiam caedit*  
 Mercurius-NOM.SG.M Sosia-ACC.SG.F frapper-3SG.PRS.IPFV.IND.ACT  
 « *Mercure frappe Sosia.* »
- b. *Sosia caeditur*  
 Sosia-NOM.SG.F frapper-3SG.PRS.IPFV.IND.PASS  
 « *Sosia est frappée.* »

Cependant, il existe d'autres moyens qui mènent à la construction de formes utilisées dans des constructions s-passives. Flobert (1967) indique en tout quatre moyens permettant d'exprimer le passif. Sur ces quatre, nous n'en retenons cependant que trois, le dernier ne correspondant pas à la définition du passif donnée ci-dessus <sup>11</sup>.

a) Il existe quelques exemples isolés où un changement de voyelle thématique (IACEO/IACIO) ou l'ajout d'un suffixe dérivationnel (AUGEO/AUGESCO) produit la forme passive correspondant à un actif (Flobert, 1967, p.574).

Les exemples donnés par Flobert sont les suivants (actif à gauche, passif à droite) : IACIO/IACEO « *reposer* », AUGEO/AUGESCO « *renforcer* », TENDO/TENEO « *dé-*

---

11. Flobert (1967) indique notamment que le moyen le plus simple d'obtenir une forme passive (c'est-à-dire une forme acceptable dans une construction passive) revient à une conversion directe de l'actif en le passif. Flobert donne aux verbes permettant une telle conversion le nom de verbes *ambivalents* ou *équivoques* : « *Enfin le passif le plus économique consiste à conserver le verbe actif en substituant un sujet d'une autre catégorie ou en modifiant la construction.* » (Flobert, 1967, p.574). Des exemples de ces verbes donnés par l'auteur sont : ABSTINEO « *tenir éloigné* », INSINUO « *tordre, faire connaître* », VERTO « *tourner* », DIFFERO « *dispenser* ». L'auteur précise que ce procédé devient notamment très productif à la basse époque. Par rapport à la définition ci-dessus, ces exemples sont néanmoins problématiques en ce qu'ils ne satisfont pas le critère du marquage morphologique obligatoire. De surcroît, Flobert donne également quelques exemples pour le français qu'on ne classerait pas spontanément comme des cas de constructions passives mais plutôt comme des instances de constructions neutres, cf. les verbes BRÛLER, CHANGER, CUIRE, ROUGIR (Flobert, 1967, p.574).

M-ACTIF							
			IND		SBJV		
			SG	PL	SG	PL	
IPFV	PRS	1	RAD1- <i>ō</i>	RAD1- <i>mus</i>	RAD1- <i>am</i>	RAD1- <i>āmus</i>	
		2	RAD1- <i>s</i>	RAD1- <i>tis</i>	RAD1- <i>ās</i>	RAD1- <i>ātis</i>	
		3	RAD1- <i>t</i>	RAD1- <i>nt</i>	RAD1- <i>at</i>	RAD1- <i>ant</i>	
	PST	1	RAD1- <i>bam</i>	RAD1- <i>bāmus</i>	RAD1- <i>rem</i>	RAD1- <i>rēmus</i>	
		2	RAD1- <i>bās</i>	RAD1- <i>bātis</i>	RAD1- <i>rēs</i>	RAD1- <i>rētis</i>	
		3	RAD1- <i>bat</i>	RAD1- <i>bant</i>	RAD1- <i>ret</i>	RAD1- <i>rent</i>	
	FUT	1	RAD1- <i>bo</i>	RAD1- <i>bimus</i>			
		2	RAD1- <i>bis</i>	RAD1- <i>bitis</i>			
		3	RAD1- <i>bit</i>	RAD1- <i>bunt</i>			
PFV	PRS	1	RAD2- <i>i</i>	RAD2- <i>imus</i>	RAD2- <i>erim</i>	RAD2- <i>erimus</i>	
		2	RAD2- <i>istī</i>	RAD2- <i>istis</i>	RAD2- <i>eris</i>	RAD2- <i>eritis</i>	
		3	RAD2- <i>it</i>	RAD2- <i>ērunt (-ēre)</i>	RAD2- <i>erit</i>	RAD2- <i>ērint</i>	
	PST	1	RAD2- <i>eram</i>	RAD2- <i>erāmus</i>	RAD2- <i>issem</i>	RAD2- <i>issēmus</i>	
		2	RAD2- <i>erās</i>	RAD2- <i>erātis</i>	RAD2- <i>issēs</i>	RAD2- <i>issētis</i>	
		3	RAD2- <i>erat</i>	RAD2- <i>erant</i>	RAD2- <i>isset</i>	RAD2- <i>issent</i>	
	FUT	1	RAD2- <i>erō</i>	RAD2- <i>erimus</i>			
		2	RAD2- <i>eris</i>	RAD2- <i>eritis</i>			
		3	RAD2- <i>erit</i>	RAD2- <i>erint</i>			
IMP	PRS	2	RAD1	RAD1- <i>te</i>			
		FUT	2	RAD1- <i>to</i>	RAD1- <i>tōte</i>		
			3	RAD1- <i>to</i>	RAD1- <i>tōte</i>		
INF	PRS		RAD1- <i>re</i>				
		PST		RAD2- <i>isse</i>			
				RAD3- <i>ūrus/-ūra/-ūrum esse</i>			
PTCP	PRS		RAD1- <i>ns, -ntis</i>				
		FUT		RAD3- <i>ūrus/-ūra/-ūrum</i>			
GRV	FUT			RAD1- <i>ndī/-ō/-um</i>			
		SUP		RAD3- <i>um/-ū</i>			

RAD1

RAD2

RAD3

TABLE 1.10 : Les désinences m-actives



M-PASSIF						
			IND		SBJV	
			SG	PL	SG	PL
IPFV	PRS	1	RAD1- <i>or</i>	RAD1- <i>mur</i>	RAD1- <i>ar</i>	RAD1- <i>āmur</i>
		2	RAD1- <i>ris (-re)</i>	RAD1- <i>minī</i>	RAD1- <i>āris (-āre)</i>	RAD1- <i>āminī</i>
		3	RAD1- <i>tur</i>	RAD1- <i>ntur</i>	RAD1- <i>atur</i>	RAD1- <i>antur</i>
	PST	1	RAD1- <i>bar</i>	RAD1- <i>bamur</i>	RAD1- <i>rer</i>	RAD1- <i>rēmur</i>
		2	RAD1- <i>bāris (-bāre)</i>	RAD1- <i>bāminī</i>	RAD1- <i>rēris (-rēre)</i>	RAD1- <i>rēminī</i>
		3	RAD1- <i>bātur</i>	RAD1- <i>bantur</i>	RAD1- <i>rētur</i>	RAD1- <i>rentur</i>
	FUT	1	RAD1- <i>bor</i>	RAD1- <i>bimur</i>		
		2	RAD1- <i>beris (-bere)</i>	RAD1- <i>biminī</i>		
		3	RAD1- <i>bitur</i>	RAD1- <i>buntur</i>		
PFV	PRS	1	RAD3- <i>us sum</i>	RAD3- <i>ī sumus</i>	RAD3- <i>us sim</i>	RAD3- <i>ī sīmus</i>
		2	RAD3- <i>us es</i>	RAD3- <i>ī estis</i>	RAD3- <i>us sīs</i>	RAD3- <i>ī sītis</i>
		3	RAD3- <i>us est</i>	RAD3- <i>ī sunt</i>	RAD3- <i>us sit</i>	RAD3- <i>ī sint</i>
	PST	1	RAD3- <i>us eram</i>	RAD3- <i>ī erāmus</i>	RAD3- <i>us essem</i>	RAD3- <i>ī essēmus</i>
		2	RAD3- <i>us erās</i>	RAD3- <i>ī erātis</i>	RAD3- <i>us essēs</i>	RAD3- <i>ī essētis</i>
		3	RAD3- <i>us erat</i>	RAD3- <i>ī erant</i>	RAD3- <i>us esset</i>	RAD3- <i>ī essent</i>
	FUT	1	RAD3- <i>us erō</i>	RAD3- <i>ī erimus</i>		
		2	RAD3- <i>us eris</i>	RAD3- <i>ī eritis</i>		
		3	RAD3- <i>us erit</i>	RAD3- <i>ī erint</i>		
IMP	PRS	2	RAD1- <i>re</i>	RAD1- <i>minī</i>		
		FUT	2	RAD1- <i>tor</i>		
	FUT	3	RAD1- <i>tor</i>	RAD1- <i>ntor</i>		
INF	PRS		RAD1- <i>(r)ī</i>			
	PST		RAD3- <i>us/-a/-um esse</i>			
PTCP	PST		RAD3- <i>us/-a/-um</i>			
GRD			RAD1- <i>ndus/-nda/-ndum</i>			

RAD1

RAD3

TABLE 1.11 : Les désinences m-passives

*ployer* », PENDO/PENDEO « *peser* »<sup>12</sup>.

Par ailleurs, le changement de voyelle thématique induit également un changement de classe flexionnelle, de sorte que nous pouvons observer pour ces verbes que le passif peut s'accompagner d'un changement de classe flexionnelle. Il s'agit donc d'une forme de passivation qui se traduit par un changement des propriétés flexionnelles d'un lexème. Or, formellement, un tel changement n'est généralement pas conçu comme étant du ressort de la flexion mais plutôt de la dérivation. Cet argument est plus fort encore pour les verbes formant explicitement leur passif avec le suffixe dérivationnel *-esc*.

b) Le réfléchi peut servir à former le s-passif.

Exemples : SE SERVARE (INF) « *être conservé* » de SERVARE « *préservé* »<sup>13</sup>.

c) Il existe également des cas que l'on pourrait qualifier de passivation supplétive. Ces exemples entrent dans la définition de Siewierska (2012) uniquement si l'on admet la supplétion comme marquage morphologique du passif. Exemples : FERIO/VAPULO « *frapper* », FACIO/FIO « *créer* », PERDO/PEREO « *détruire* »<sup>14</sup> (Flobert, 1967, p.573-4).

Le problème de cette troisième stratégie réside en le fait qu'il s'agit par ailleurs pour ces verbes d'entrées lexicales distinctes, les verbes servant de passif supplétif ayant eux-mêmes leur propre paradigme et un sens différent. Gaffiot & Flobert (2000) citent ainsi deux entrées séparées pour FIO, l'une étant le passif de FACIO, l'autre une entrée à part, en outre avec un passif archaïque (*fitur, fiebantur, ...*). Ainsi, ces verbes utilisés pour les constructions passives s'apparentent plus à une stratégie de réparation pour combler les trous apparaissant dans les paradigmes défectifs de verbes comme FERIO, FACIO et PERDO.

Par un renversement de point de vue, ces verbes sont par ailleurs également parfois

12. Toutefois, les données de Gaffiot & Flobert (2000) (confirmées également par Prior & Wohlberg (1995) pour les verbes répertoriés par ce dernier) montrent que ces formations du passif sont en général en concurrence directe avec une formation morphologique standard du passif par l'emploi des désinences passives. En revanche, le verbe correspondant au passif dans les couples de verbes ci-dessus ne comporte généralement pas de formes m-passives. Parmi les verbes cités ci-dessus ceux qui possèdent des formes m-passives selon Gaffiot & Flobert (2000) sont les suivants : IACIO/IACEO : IACIO/IACIOR, IACEO/- ; AUGEO/AUGESCO : AUGEO/AUGEOR, AUGESCO/- ; TENDO/TENDEO : TENDO/TENDOR, TENEO/TENEOR (défectif au parfait) ; PENDO/PENDEO : PENDO/PENDOR, PENDEO/-.

13. Flobert (1967) donne également l'exemple inverse du m-passif servant à exprimer une valeur réfléchie : TRADOR « *se rendre* » de TRADO « *abandonner* ».

14. VAPULO « *recevoir des coups* » et PEREO « *disparaître, périr* » sont également des verbes indépendants intransitifs. En revanche, ils ne possèdent à ce titre pas de formes m-passives. FIO « *arriver, devenir* » possède guère qu'un m-passif archaïque défectif.

qualifiés d'*activa tantum* (Kiparsky, 2005), dénotant par là leur qualité de verbes transitifs défectifs auxquels il manque les formes passives attendues. Si l'on adopte ce point de vue, le passif en latin ne relève (au mieux) pas de la flexion canonique au sens de Corbett (2007a). Selon Corbett, la flexion ne comporte canoniquement pas de trous arbitraires dans le paradigme ; en d'autres termes, toute forme qui devrait appartenir au paradigme d'un lexème est canoniquement produite. Les cas où il y a des formes manquantes sont des cas de défectivité qui s'éloignent du canon attendu pour la flexion. En revanche, pour Corbett, ces mêmes trous arbitraires constituent bien une propriété canonique de la dérivation : pour un procédé dérivationnel donné s'appliquant à une classe de lexèmes donnée, il est canonique qu'il y ait dans cette classe des exceptions lexicales arbitraires pour lesquelles le procédé dérivationnel ne s'applique pas. En ce sens, le comportement des verbes dits *activa tantum* s'apparente bien plus à un comportement canoniquement dérivationnel que flexionnel.

### 1.1.2.2 La valeur du passif morphologique (m-passif)

En ce qui concerne la valeur associée au m-passif, le cas habituel est de considérer que le m-passif sert à marquer le s-passif. Toutefois, Flobert (1967) cite cinq valeurs concurrentes à l'expression du s-passif qui chacune concerne un nombre significatif de lexèmes<sup>15</sup>.

- a) Il existe une valeur médio-passive des m-passifs qui permet d'indiquer la dimension volontaire associée à une action :

*a se moveri* (de soi mouvoir.M-PASSIF.INF) « *se mouvoir de son propre fait* »  
*tundis palmis, plangi alis* (M-PASSIF.INF) « *se frapper de ses mains (ailes)* »

- b) Les verbes<sup>16</sup>, indiquent également un changement de place (mouvement) ou de situation (transformation) au m-passif :

Mouvement : DĒICIOR « *évacuer* », FEROR « *déplacer* », MOVEOR « *s'agiter* », VEHOR « *se déplacer à/en X* », VOLVOR « *rouler* ».

Transformation : STERNOR « *s'allonger* », SOCIOR « *(se) cicatriser* », TRĀDOR « *transmettre* », VIDEOR « *apparaître* », CRUCIOR « *souffrir* », DĒLECTOR « *se réjouir* », ALOR « *se nourrir (de)* », AMĪCOR « *agir en ami* », LAVARI « *se baigner* ».

Flobert indique avoir compté plus de 700 verbes pouvant avoir cette valeur.

- c) Un autre exemple, lexicalement contraint, concerne les verbes de mouvement à dimension volitive.

---

15. Les exemples de verbes sont ceux de Flobert (1967) ; les traductions sont celles indiquées pour les verbes concernés par Gaffiot & Flobert (2000) et Lewis & Short (1879).

16. Les verbes *mutatifs* chez Flobert (1967).

*se (ipsum) movere* (soi (soi-même) mouvoir.M-ACTIF.INF) « *se mettre en mouvement volontairement* »

MOVERI (mouvoir.M-PASSIF.INF) « *être en mouvement, se mouvoir* »

Exceptionnellement seulement : MOVERI (AB) ABL (mouvoir.M-PASSIF.INF (par) qqn) « *être mis en mouvement par quelqu'un* » (s-passif)

Autres exemples : AGITOR « *se mettre en mouvement, s'agiter* », CIEOR « *s'émouvoir, se mettre en mouvement* », IMPELLOR « *mettre en mouvement* ».

Cette classe de verbes comporte au moins 500 lexèmes selon Flobert.

d) Les passifs impersonnels (au moins 400 lexèmes selon les comptes de Flobert (1967)) :

*agitur* « *ça va* »

*fertur* « *on rapporte (que)* »

*itur* « *on va* »

*vivitur* « *on vit* »

Au cours de l'évolution de la langue, ce phénomène s'est ensuite étendu à d'autres unités comme *clamatur*, *concurritur*, *pugnatur*, et de même au perfectif correspondant *clamatum est*, ....

Flobert (1967) indique enfin que le passif n'est pas une propriété exclusive des verbes transitifs en latin. En effet, avec le développement du passif impersonnel, les formes de passifs intransitifs se sont font de plus en plus nombreuses. Ce développement semble néanmoins se restreindre à des verbes employés dans la langue administrative, religieuse ou dans des expressions de « *prudence paysanne* » (Flobert, 1967, p.578), ou de l'annalistique. La capacité à produire un passif impersonnel apparaît donc moins comme une propriété liée à la transitivité du verbe en question que comme une propriété lexicale de « *verbes susceptibles d'exprimer un procès anonyme ou collectif* » (Flobert, 1967, p.578).

### 1.1.2.3 Conclusion sur la valeur du passif en latin

Les exemples exposés dans les sections 1.1.2.1 à 1.1.2.2 montrent que le m-passif latin n'est pas nécessairement le marquage du passif tel que défini plus haut. En particulier, il n'exprime pas un alignement unique entre rôles thématiques et structure argumentale. D'autres valeurs, telles le réfléchi, la volition ou l'impersonnel peuvent être exprimées par le m-passif latin. Nous récapitulons le panorama présenté par Flobert dans le tableau 1.12 indiquant en colonne le type de construction et en ligne le marquage morphologique associé.

VALEUR	M-ACTIF	M-PASSIF
S-ACTIF	<i>standard</i>	<i>verbes déponents</i>
S-PASSIF	<i>+se</i> <i>oui</i> <i>cstr. neutre</i>	<i>standard</i>
RÉCIPROQUE	<i>oui</i>	<i>oui</i>
RÉFLÉCHI	<i>oui</i>	<i>oui</i>

TABLE 1.12 : Constructions valentielles disponibles avec un marquage m-actif/m-passif selon Flobert (1967)

D'après les données (notamment diachroniques) relevées par Flobert (1967), nous observons également que de façon de plus en plus manifeste à travers le temps le m-passif semble évoluer en un marquage lexicalement déterminé mais décorrélé de la voix en tant que telle. En d'autres termes, le m-passif acquiert les propriétés attendues d'une classe flexionnelle indépendante.

## 1.2 Deux analyses contemporaines des déponents latins

Notre présentation des données s'est essentiellement nourrie de la description de Flobert (1967). Cette riche description donne notamment des indications détaillées sur les valeurs des formes verbales et sur l'étendue des phénomènes observables. Ce travail constitue une source précieuse sur le fonctionnement de la flexion verbale latine et en particulier sur le comportement des verbes déponents par rapport au reste du système linguistique. L'auteur accompagne notamment son étude d'indications quantitatives permettant de mesurer l'ampleur des phénomènes décrits.

Dans cette section nous nous intéressons à deux analyses contemporaines des verbes déponents : Kiparsky (2005) et Hippisley (2007). Nous observerons en particulier la cohérence interne des deux modèles proposés et leur adéquation aux observations linguistiques présentées dans la section 1.1. La comparaison de ces deux analyses nous servira pour illustrer les caractéristiques méthodologiques que nous cherchons à mettre en place pour la formalisation de descriptions et d'analyses morphologiques. Nous en proposerons un exemple par notre propre analyse à la section 1.4.

## 1.2.1 Analyse flexionnelle des déponents latins : Hipsisley (2007)

### 1.2.1.1 L'analyse de Hipsisley (2007)

L'analyse de Hipsisley s'inscrit dans la famille des approches lexématiques de type inférentiel-réalisationnel (Stump, 2001). L'unité de base du lexique est le lexème. Chaque lexème est associé à une classe flexionnelle définissant la réalisation de ses formes. L'analyse de Hipsisley en *Network Morphology* (Corbett & Fraser, 1993 ; Brown & Hipsisley, 2012) suit la tradition des approches flexionnelles du passif. Cette analyse du comportement des déponents latins est sans doute la plus aboutie d'une série d'approches comportant notamment celles de Stump (2002, 2006) et de Sadler & Spencer (2001). L'analyse est accompagnée d'une implémentation explicite en DATR (Evans & Gazdar, 1996)<sup>17</sup>, consultable en ligne<sup>18</sup>. Par son implémentation complète, elle présente l'avantage de traiter et de générer l'intégralité des formes des verbes étudiés.

Dans l'analyse de Hipsisley, m-actif et m-passif constituent des sous-paradigmes distincts du paradigme complet d'un verbe transitif. M-actif et m-passif y sont par défaut alignés avec les s-actif et le s-passif respectivement. La passivation se traduit par un changement d'exposant au sein du même paradigme, illustré par la figure 1.1.

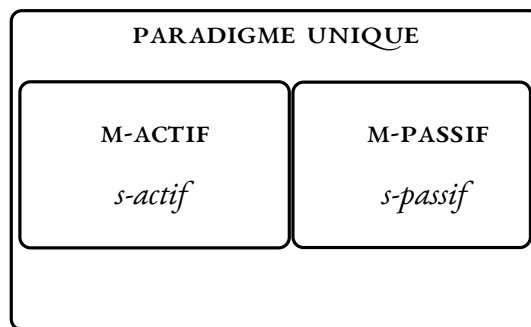


FIGURE 1.1 : Le m-passif en tant que sous-paradigme des paradigmes de verbes transitifs

L'implémentation DATR de l'analyse de Hipsisley repose sur une hiérarchie de nœuds qui héritent les uns des autres. Dans l'implémentation en DATR, chaque entrée lexicale verbale est associée à son propre nœud (cf. figure 1.2 page 53) et hérite par défaut<sup>19</sup> d'un nœud supérieur appelé VERB (cf. figure 1.3 page 53). Le nœud VERB précise ce qui est commun à tous les verbes, en l'occurrence d'avoir une morphologie (mor) qui produit

17. Voir le chapitre 2 pour une description de *Network Morphology*.

18. Sur le site associé au livre *Network Morphology : a Defaults-based Theory of Word Structure* (Brown & Hipsisley, 2012) : [http://networkmorphology.as.uky.edu/sites/default/files/ch5\\_latin.dtr](http://networkmorphology.as.uky.edu/sites/default/files/ch5_latin.dtr) (valable au 11 juin 2013).

19. L'héritage est représenté par <> == .

des formes actives et passives tel que spécifié par les nœuds `ACT_FORMS` et `PASS_FORMS` correspondants (cf. figure 1.4 page 53). Les nœuds pour les formes actives précisent à leur tour des nœuds permettant d'obtenir l'information nécessaire à la construction des formes imperfectives, d'une part, et perfectives de l'autre. L'architecture globale peut ainsi être vue comme un graphe orienté dont nous illustrons un extrait dans la figure 1.8 page 56.

Les nœuds `ACT_FORMS` et `PASS_FORMS` représentent donc le point de départ des chemins permettant de construire le sous-paradigme actif d'une part, et le sous-paradigme passif de l'autre.

Hippisley représente les verbes déponents comme des verbes qui par défaut se comportent comme des verbes suivant les indications du nœud `VERB`. Mais pour les formes s-actives, au lieu de suivre les indications générales stipulées pour les verbes dans le nœud `VERB`, ils suivent un schéma distinct, défini dans la figure 1.5 page 54. La règle de renvoi ligne 3 indique que les formes s-actives doivent être construites comme des formes m-passives, à savoir en suivant le chemin indiqué ligne 4 dans la figure 1.3 page 53 pour le nœud `VERB`. La modélisation est donc analogue à celle d'un phénomène de synchrétisme directionnel, et consiste à emprunter l'intégralité du sous-paradigme m-passif pour exprimer le s-actif.

La seule exception à la règle de renvoi ligne 3 est indiquée ligne 4 : pour l'infinitif futur imperfectif, forme inexistante dans le sous-paradigme m-passif, les verbes déponents suivent exceptionnellement le défaut `VERB` au lieu de suivre leur schéma particulier `DEPONENT`. Enfin, le nœud `DEPONENT` spécifie explicitement (ligne 5) que le s-passif n'est pas défini pour les verbes déponents. On peut noter que le traitement est analogue à celui d'un verbe défectif.

Les verbes déponents mixtes suivent respectivement les nœuds `PERFECT_DEPONENT` et `IMPF_DEPONENT` illustrés par la figure 1.6 page 54. Ils se comportent par défaut comme les verbes déponents et par conséquent comme indiqué par le nœud `DEPONENT`. En revanche les DM1 se distinguent des verbes déponents en suivant le schéma par défaut de `VERB` à l'imperfectif (cf. ligne 3), tandis que les DM2 suivent reviennent au défaut de `VERB` pour le seul perfectif (cf. ligne 7).

Les entrées lexicales associées pour des verbes déponents et des verbes déponents mixtes sont indiquées dans la figure 1.7 page 55.

Dans la figure 1.8 page 56 récapitulant graphiquement l'analyse de Hippisley, chaque entrée héritant directement du nœud `VERB` est considérée comme ayant une réalisation pour le s-actif et le s-passif selon les propriétés définies dans le graphe contenant `ACT_FORMS` et `PASS_FORMS`. Les alignements entre syntaxe et morphologie par défaut

```
1
2 Amo:
3   <> == VERB
4   <gloss> == love
5   <root> == am
6   <stem> == CONJ_1.
7
8
9 Monêo:
10  <> == VERB
11  <gloss> == advise
12  <root> == mon
13  <stem> == CONJ_2.
14
15 Rego:
16  <> == VERB
17  <gloss> == rule
18  <root> == reg
19  <stem> == CONJ_3.
20
21 Audio:
22  <> == VERB
23  <gloss> == hear
24  <root> == aud
25  <stem> == CONJ_4.
```

FIGURE 1.2 : Définition des entrées lexicales pour les quatre types de conjugaison

```
1 VERB:
2   <syn> == "<mor>"
3   <mor active> == ACT_FORMS:<>
4   <mor passive> == PASS_FORMS:<>.
```

FIGURE 1.3 : Définition de la morphologie verbale

```
1 ACT_FORMS:
2   <imperfective> == ACT_IMPF:<>
3   <perfect> == ACT_PERF:<>.
4
5 PASS_FORMS:
6   <imperfective> == PASS_IMPF:<>
7   <perfect> == PASS_PERF:<>.
```

FIGURE 1.4 : Définition des formes verbales



```

1 DEPONENT :
2   <> == VERB
3   <syn active> == "<mor passive>"
4   <syn active imperfective future infinitive> == VERB
5   <syn passive> == undefined.

```

FIGURE 1.5 : Représentation de la déponence

```

1 PERFECT_DEPONENT :
2   <> == DEPONENT
3   <syn active imperfective> == VERB.
4
5 IMPF_DEPONENT :
6   <> == DEPONENT
7   <syn active perfect> == VERB.

```

FIGURE 1.6 : Règles de renvoi spécifiques pour les semi-déponents (verbes déponents mixtes)

sont représentés pas les flèches vertes. Les verbes déponents, associés au nœud `DEPONENT`, affichent un alignement déviant, représenté par les flèches rouges. L'alignement exprimé par les flèches rouges remplace l'alignement par défaut des flèches vertes pour toutes les entrées héritant directement du nœud `DEPONENT`. Les verbes déponents mixtes quant à eux affichent un alignement différent des verbes déponents classiques. Cet alignement est représenté par les flèches bleues. Les flèches bleues spécifient en particulier un retour partiel au cas général, à savoir que l'exception formulées par les flèches rouges est annulée au profit du cas par défaut, représenté par les flèches vertes partant du nœud `VERB`. Pour les cas non explicité par les flèches bleues, l'alignement exprimé par les flèches rouges est valable par défaut.

Ainsi, Hippius propose une analyse qui présente implicitement les verbes déponents comme des verbes défectifs à qui il manque les formes du sous-paradigme m-actif pour l'expression du s-actif. L'auteur en déduit une représentation des verbes déponents comme un cas de syncrétisme directionnel qui provoque l'emploi syncrétique de formes m-passives pour l'expression du s-actif<sup>20</sup>, comme si, en l'absence de formes m-actives disponibles, le s-actif *empruntait* les formes du s-passif pour être réalisé. L'analyse syncrétique des verbes déponents est illustrée par la figure 1.9 page 57.

Comme indiqué dans la présentation des données ci-dessus, on trouve toutefois également quelques formes actives dans les paradigmes des verbes déponents. Hippius ex-

20. Représenté par la règle `<syn active> == "<mor passive>"`.

```

1 Hortor:
2   <> == DEPONENT
3   <gloss> == encourage
4   <root> == hort
5   <stem> == CONJ_1.
6
7 Uutor:
8   <> == DEPONENT
9   <gloss> == use
10  <root> == ūt
11  <stem 3> == ūs
12  <stem> == CONJ_3.
13
14 Audeo:
15  <> == PERFECT_DEPONENT
16  <gloss> == dare
17  <root> == aud
18  <stem 3> == aus
19  <stem> == CONJ_2.
20
21 Revertor:
22  <> == IMPF_DEPONENT
23  <gloss> == return
24  <root> == revert
25  <stem 2> == <root>
26  <stem 3> == <root> s
27  <stem> == CONJ_3.

```

FIGURE 1.7 : Entrées lexicales pour les déponents (verbes déponents) et semi-déponents (verbes déponents mixtes)

plique leur présence de la façon suivante. Les formes m-actives n'ont pas de correspondance dans le paradigme m-passif. L'emprunt du paradigme m-passif n'affecte donc pas les cases concernées du paradigme, qui restent de ce fait exprimées par des formes m-actives en vertu du défaut hérité du nœud VERB<sup>21</sup>.

Le même mécanisme permet également d'expliquer la structure des paradigmes des verbes déponents mixtes. Pour les les verbes déponents mixtes de type 1, les DM1, qui affichent des formes m-actives à l'imperfectif et des formes m-passives au perfectif, il suffit de considérer que les DM1 possèdent un sous-paradigme m-actif défectif au seul perfectif; ce sont donc ces seules formes qui seront empruntées au sous-paradigme m-passif à l'aide

21. L'emprunt des formes m-passives pour exprimer le s-actif est représenté par des règles de renvoi en DATR. Ainsi, l'emprunt des formes m-passives est effectué globalement pour le sous-paradigme passif. Ce fait explique notamment la présence de participes passés dans le paradigme d'un verbe déponent, forme pourtant absente des paradigmes de verbes m-actifs. En revanche, même lorsqu'une forme n'existe pas dans le sous-paradigme passif, comme c'est par exemple le cas pour les participes actifs et le gérondif, la forme m-active est employée par défaut dès lors qu'elle existe dans le paradigme du verbe déponent en question.

FIGURE 1.8 : Représentation sous forme de graphe de l'analyse de Hippisley (2007)

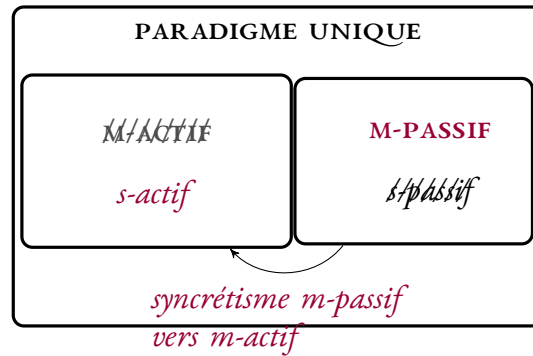


FIGURE 1.9 : Modélisation des paradigmes des verbes déponents par syncrétisme

du syncrétisme directionnel décrit ci-dessus. Pour les DM2, le même procédé s'applique, avec les formes imperfectives cette fois.

L'analyse de Hippiisley rend donc élégamment compte des formes surnuméraires des verbes déponents par rapport à un paradigme m-passif standard.

Pour les verbes déponents transitifs, ce phénomène de syncrétisme a par ailleurs pour conséquence qu'il leur manque la moitié de leur paradigme. Les formes une fois empruntées pour exprimer le s-actif ne sont plus disponibles pour exprimer le s-passif<sup>22</sup>. L'indisponibilité du s-passif doit être stipulée indépendamment<sup>23</sup>. La modélisation de Hippiisley rend donc bien compte du point c) relevé par Flobert (1967) à propos des verbes déponents : Il y a des verbes déponents transitifs qui se prêtent à la passivation, mais leur paradigme passif est dans ce cas en général défectif, souvent réduit au seul participe passé passif.

Ainsi, Hippiisley résume le phénomène observable pour les verbes déponents (« *déponents* ») de la façon suivante :

- a) Il y a un décalage entre contenu (actif) et forme (passive).
- b) Les verbes déponents ont un paradigme de formes passives qui contient néanmoins quelques formes actives.
- c) Ils correspondent également à un paradigme de contenu actif, ce dernier ne comprenant habituellement pas de participe passé ; pourtant les déponents en possèdent.
- d) Il existe des semi-déponents dont les paradigmes contiennent à la fois des formes actives et passives alors même que leur paradigme de contenu est

22. Le fait que les formes m-passives ne soient plus disponibles pour l'expression du s-passif correspond à la définition même de la déponence par Corbett (2007b).

23. En l'occurrence par la règle <syn passive> == undefined.

uniformément actif.

- e) Les semi-déponents peuvent afficher des distributions variées de formes actives et passives, distribuées librement selon un découpage aspectuel ; ce découpage est donc variable au sein de la classe des semi-déponents.

### 1.2.1.2 Discussion

L'analyse de Hippiisley (2007) est entièrement formalisée et implémentée en DATR. Elle génère (au moins pour tous les verbes contenus dans l'implémentation), l'ensemble des formes correctes.

1. Cependant, faute de prendre en compte un lexique suffisamment étendu, les données de verbes non standard ne sont pas prises en compte. Les verbes passivés dont la sémantique et la valeur syntaxique diffèrent de la simple expression du s-passif ne sont pas étudiés en tant que tels par Hippiisley (2007). Pourtant, comme nous l'avons vu à la section 1.1.2.2, ils ne constituent pas quelques cas isolés, mais bien un sous-ensemble représentatif du lexique verbal latin<sup>24</sup>. Aucun alignement ne permet notamment d'illustrer l'emploi du m-passif pour exprimer les valeurs de s-passif, médio-passif, « *mutatifs* », mouvement en cours et d'impersonnel.
2. Une telle imprédictibilité du changement sémantique n'est en général pas associée à une opération flexionnelle, mais plutôt à une opération dérivationnelle. Notamment Wurzel (1989, p. 36)<sup>25</sup>, citant Kubrjakova (1972) et Isačenko (1962), indique :

« *In inflection, the difference in meaning between basic and derived forms is fully predictable, whereas in derivation it is only partially so.* »

« *En flexion, la différence de sens entre le sens de base<sup>26</sup> et le dérivé<sup>27</sup> est entièrement prédictible, tandis qu'en dérivation elle ne l'est que partiellement.* »

Selon cette définition, le passif en latin devrait donc plutôt être conçu comme un phénomène dérivationnel. Le principal défaut de l'analyse de Hippiisley semble donc résider dans le fait de ne pas poser la question théorique préalable à l'analyse concernant la nature de l'opération du passif qui découle justement des données non couvertes.

3. L'origine même des formes passives empruntées pose problème dans le cas des verbes déponents intransitifs. En effet, en tant que verbes intransitifs, leur paradigme ne

---

24. Voir les comptes indiqués par Flobert (1967), cf. section 1.1.2.2.

25. Voir aussi Boyé (2011).

26. Comprendre « *le sens inhérent au lexème* ».

27. Comprendre « *la forme fléchie* ».

devrait pas contenir de formes m-passives et il semble par conséquent difficile de les emprunter, fût-ce pour un emploi s-actif<sup>28</sup>.

S'ajoute à ces problèmes de fond un problème, mineur, de modélisation. En effet, comme nous l'avons mentionné plus haut, l'auteur indique qu'après l'application du syncrétisme directionnel aux formes passives<sup>29</sup>, les formes m-passives ne sont plus disponibles, notamment pour former le sous-paradigme passif attendu pour des verbes déponents transitifs. En cela, le modèle prédit bien la défektivité au passif des verbes déponents transitifs.

Cependant, on ne voit pas très bien comment le modèle proposé par Hippiisley distingue les cas de déponence, cas qu'il illustre par les verbes déponents, des cas de syncrétismes simples.

Considérons par exemple le syncrétisme entre le nominatif et l'accusatif des noms neutres en latin, cf. tableau 1.13. En *Network Morphology*, il sera modélisé par le même mécanisme formel de règles de renvoi que celui employé pour rendre compte des verbes déponents. Pourtant, dans ce cas, malgré le syncrétisme de l'accusatif vers le nominatif<sup>30</sup>, la forme de l'accusatif reste disponible pour l'accusatif. Le paradigme des noms neutres en latin n'est pas défectif.

Or, cette différence entre formes disponibles et indisponibles après l'application des règles de renvoi est justement la différence majeure que Corbett (2007b) note entre la déponence et le syncrétisme. En simplifié, il décrit le syncrétisme comme une opération de *copier-coller*, tandis que la déponence correspond à du *couper-coller*. Dans la modélisation de Hippiisley, rien ne permet d'expliquer pourquoi les formes m-passives ne sont de fait plus disponibles pour former du passif de verbes déponents transitifs. Tout ce qu'y est indiqué est la stipulation indépendante <syn passive> == undefined.

---

28. À moins d'analyser l'intégralité des verbes intransitifs comme des verbes défectifs au passif et de considérer le paradigme à 143 cases comme seul paradigme régulier et complet pour les verbes latins. Une telle approche serait cependant à la fois contre-intuitive (il ne semble en effet pas y avoir de formes passives manquantes dans les paradigmes des verbes intransitifs), compliquée par l'introduction importante de paradigmes irréguliers qu'elle supposerait et par conséquent hautement artificielle.

29. Après l'application des règles des renvoi en *Network Morphology*.

30. Nous ne justifions pas ici le choix de la direction du syncrétisme. Pour plus d'informations à ce sujet, voir les chapitres 4 et 7. Ce syncrétisme n'est d'ailleurs pas le seul dans le paradigme représenté par le tableau 1.13.

	SG	PL
NOM	<i>bell-um</i>	<i>bell-a</i>
ACC	<i>bell-um</i>	<i>bell-a</i>
GEN	<i>bell-i</i>	<i>bell-orum</i>
DAT	<i>bell-o</i>	<i>bell-is</i>
ABL	<i>bell-o</i>	<i>bell-is</i>

TABLE 1.13 : Syncrétismes dans le paradigme d'un nom neutre en latin (BELLUM « la guerre »)

## 1.2.2 Une approche lexicale des déponents latins : Kiparsky (2005)

### 1.2.2.1 Modélisation des verbes déponents selon Kiparsky

Kiparsky (2005) fonde son analyse sur un inventaire préalable des valeurs du passif qui reprend – avec certes moins de détails — les données qui avaient été présentées par Flobert (1967, 1975) avant lui. Kiparsky note ainsi en introduction de son analyse que le passif morphologique en latin peut correspondre à plusieurs valeurs distinctes, toutes plus ou moins clairement liées à de la *détransitivisation*.

« *It is well-known that passive morphology in Latin has several distinct functions, most but not all of them detransitivising in some sense.* »

« *Il est bien connu que la morphologie passive en latin peut avoir plusieurs fonctions distinctes, dont principalement, mais pas uniquement, celle de détransitiviser.* »

(Kiparsky, 2005, p.7)

Suite à son inventaire empirique des valeurs du m-passif en latin décrit ci-dessus, Kiparsky (2005) propose l'analyse suivante pour rendre compte des verbes déponents.

En un premier temps, Kiparsky définit ce qu'il appelle la *flexion passive*<sup>31</sup>.

« *Passive inflection in Latin is a conjugational feature — we'll call it [ $\pm$ PASSIVE] — which can be lexically specified, for verb stems as well as for inflectional endings, or left unspecified. This feature classifies stems and endings into three types each.* »

« *La flexion passive en latin est un trait conjugationnel — nous le nommerons désormais [ $\pm$ PASSIF] — qui peut être spécifié (ou laissé sous-spécifié) dans le lexique*

31. Depuis une approche *Mot et Paradigme*, le terme de « flexion passive » ainsi que celui de « trait conjugationnel » suggèrent *a priori* l'existence d'une classe flexionnelle du passif qui serait cohérente avec les observations sur la valeur du passif faites par Kiparsky. Cependant, la notion de classe flexionnelle ne fait pas en tant que telle partie du cadre théorique dans lequel se situe Kiparsky.

*pour les radicaux verbaux comme pour les désinences flexionnelles. Ce trait permet de déterminer trois types respectivement de radicaux et de désinences. »*

Les trois types de radicaux et de désinences en question sont ainsi définis comme suit :

a) **Radicaux verbaux :**

- a) [ $\pm$ PASSIF] : radicaux qui forment des formes actives autant que passives, cf. laud-o « louer ».
- b) [+PASSIF] : radicaux des verbes déponents, cf. hort-or « exhorter ».
- c) [-PASSIF] : activa tantum, cf. perd-o « périr ».

b) **Désinences :**

- a) [ $\pm$ PASSIF] : désinences indifférenciées, cf. PTCP.PRS -ns (/nt-is/)
- b) [+PASSIF] : désinences passives, cf. 3SG.PRS.IND.PASS -tur
- c) [-PASSIF] : désinences actives, cf. 3SG. PRS.IND.ACT -t

Dans le modèle de Kiparsky (2005) ces radicaux et désinences sont stockés individuellement dans le lexique.

Kiparsky indique que la plupart des verbes sont sous-spécifiés et portent la valeur [ $\pm$ PASSIF]. Ils peuvent par conséquent se combiner avec tous les types de désinences.

La figure 1.10 résume graphiquement l'analyse de Kiparsky (2005). Les flèches représentent les combinaisons licites entre radicaux et désinences selon l'inventaire de traits spécifié par l'auteur.

### 1.2.2.2 Discussion

Dans l'approche formelle qu'esquisse Kiparsky, le système de traits [ $\pm$ PASSIF] permet de prédire correctement toutes les formes des verbes transitifs, intransitifs, des activa tantum et de différents types de verbes déponents et de verbes déponents mixtes. Toutefois, compte tenu du cadre théorique dans lequel se situe Kiparsky, la caractérisation de l'opération morphologique permettant de construire les formes passives reste non spécifiée dans son



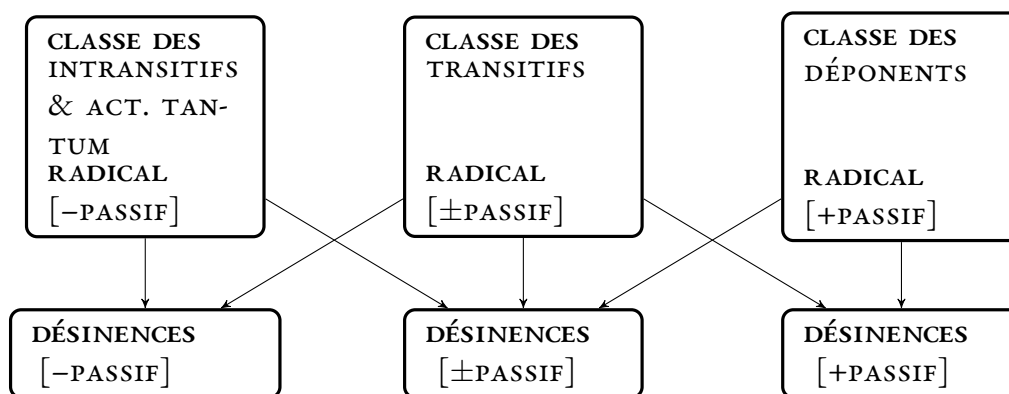


FIGURE 1.10 : Traits conjugationnels des radicaux et des désinences

analyse — malgré les constats sur la valeur du passif placé en introduction de son analyse par l'auteur.

Dans un cadre lexématique, en revanche, cette distinction serait cruciale. L'analyse de Kiparsky n'est transposable dans un cadre lexématique que si l'on suppose l'existence de paradigmes de verbes transitifs qui contiendraient à la fois des formes m-actives et des formes m-passives. Formellement, la passivation ressemblerait alors plus à une opération flexionnelle que dérivationnelle. L'analyse ne rendrait alors justement pas compte du caractère dérivationnel du passif qu'a pourtant mis en évidence Kiparsky par la présentation préalable des données : le passage aux formes passives s'accompagne de changements sémantiques relativement imprévisibles et ressemble donc plus à une opération dérivationnelle que flexionnelle.

Peut-on néanmoins adapter le modèle de Kiparsky afin qu'il rende compte des caractéristiques dérivationnelles du passif ? Dans le modèle de Kiparsky les radicaux et les désinences sont stockés indifféremment dans le lexique. Une approche dérivationnelle avec une telle vision du lexique pourrait ainsi spécifier que le passif permet de construire un nouveau radical [+passif] à partir d'un radical [-passif]. Ainsi tous les radicaux seraient spécifiés par rapport au trait [±PASSIF], seules certaines désinences seraient encore sous-spécifiées.

L'opération de dérivation pourrait ainsi être représentée comme suit.

1.4	AMO	<i>passivation</i>	AMOR
	[- <i>passif</i> ]	→	[+ <i>passif</i> ]
	« <i>aimer</i> »		« <i>être aimé (par quelqu'un)</i> »

Dans cette adaptation du modèle nous considérons que le trait  $[\pm\text{PASSIF}]$  correspond à une caractéristique d'entrées lexicales verbales (les radicaux) d'une part et à un indice de classe flexionnelle de l'autre. Ainsi, le choix d'un ensemble de désinences  $[\text{+PASSIF}]$ ,  $[\text{-PASSIF}]$  ou  $[\pm\text{PASSIF}]$ , s'apparente à l'appartenance d'un lexème à une certaine classe flexionnelle, les désinences sous-spécifiées étant communes aux deux classes.

La figure 1.11 illustre une telle analyse dérivationnelle des passifs. Le passif y correspond à un changement de classe flexionnelle. Dans la représentation donnée, la classe flexionnelle A comporte les désinences à valeurs  $[\text{-PASSIF}]$  ainsi que les désinences sous-spécifiées ( $[\pm\text{PASSIF}]$ ), tandis que la classe flexionnelle B comporte les désinences  $[\text{+PASSIF}]$  et également les désinences  $[\pm\text{PASSIF}]$ .

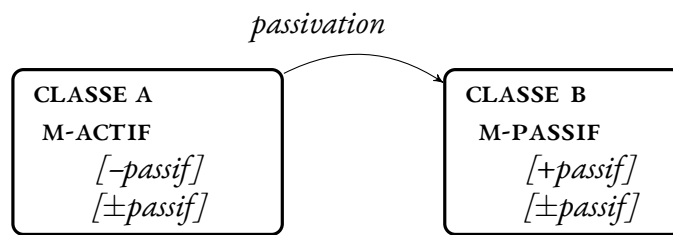


FIGURE 1.11 : La passivation comme opération dérivationnelle

Une telle adaptation de l'analyse de Kiparsky nous laisse enfin avec un modèle correspondant à l'illustration proposée dans la figure 1.12.

Dans notre adaptation de l'analyse de Kiparsky, les verbes transitifs standard de la classe A ont un correspondant intransitif appartenant à la classe B. Plus précisément, ils sont représentés par deux radicaux distincts, dérivables entre eux, la passivation étant une opération de détransitivisation. Les autres verbes appartiennent à une classe flexionnelle unique : les intransitifs standard appartiennent à la classe  $[\text{-PASSIF}]$ , tout comme les activa tantum ; les verbes déponents sont, eux,  $[\text{+PASSIF}]$ .

Enfin, en ce qui concerne les verbes déponents mixtes, Kiparsky indique :

*« the class of semi-deponent verbs (such as gaudet, gāvīsus est 'rejoices', audet, ausus est 'dares') has Perfect stems which must be listed as inherently passive. »*

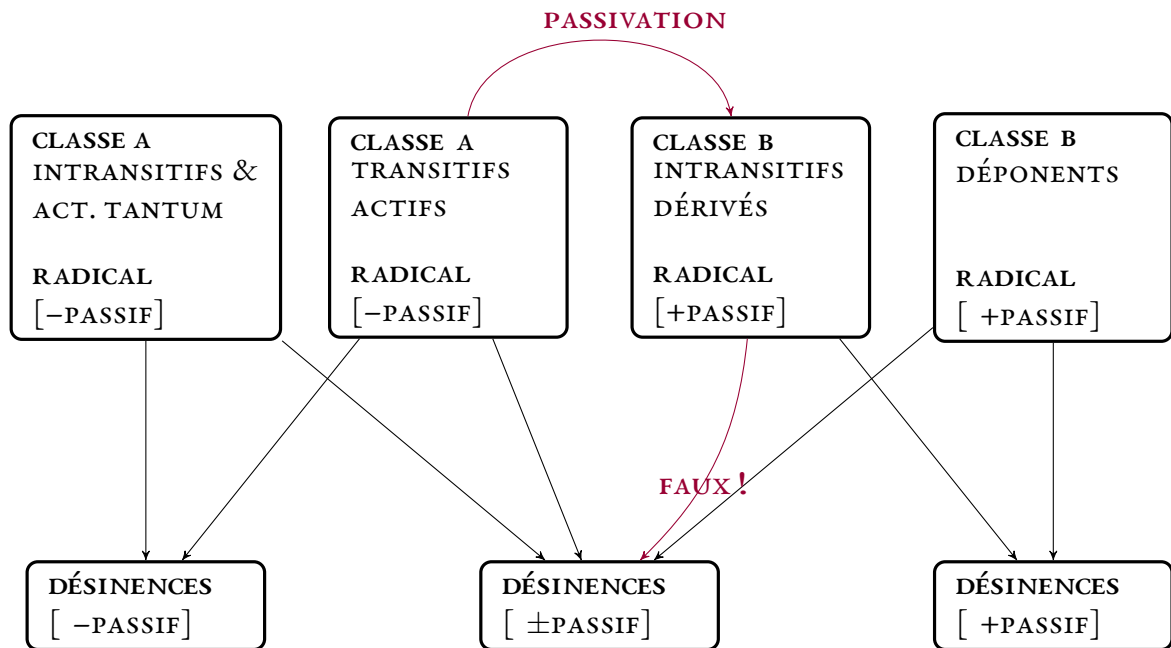


FIGURE 1.12 : Classes flexionnelles et désinences possibles

« la classe des verbes déponents (comme *gaudet*, *gāvīsus est*, « se réjouir », ou *audet*, *ausus est*, « oser ») a des radicaux du parfait qui doivent être listés dans le lexique comme étant passifs de façon inhérente. »

Les verbes déponents mixtes possèdent ainsi deux radicaux distincts. Ils utilisent par ailleurs, avec chacun de leurs deux radicaux, une partie de paradigme formée à partir de la classe appropriée. Ils peuvent ainsi être considérés comme *hétéroclites* du fait de leur paradigme mixte et *supplétifs* du fait de leur second radical.

Le modèle que nous avons suggéré pour rendre compte du caractère dérivationnel de la passivation a cependant un défaut majeur. Il surgénère des formes pour les verbes passifs standard. Il prédit en effet que tous les verbes, qu'ils soient actifs, passifs, verbes déponents ou verbes déponents mixtes, peuvent se combiner avec les désinences sous-spécifiées. Ces désinences font ainsi partie des deux classes flexionnelles *disponibles*. Pourtant les verbes passifs ne devraient pas avoir accès à ces désinences. La forme *amans* « aimant » appartient au paradigme du verbe actif *amo* « aimer » et non à celui du verbe passif dérivé *amor* « être aimé ». Il est pourtant prédit comme une forme possible des deux paradigmes. La modification de l'analyse de Kiparsky (2005) ne permet donc pas de rendre compte de façon satisfaisante des formes participiales actives que l'on observe dans les paradigmes de verbes déponents mais pas dans ceux des verbes passifs standard.

En somme, située dans un cadre théorique ne permettant pas de rendre compte

par des opérations exclusivement morphologiques de la flexion et de la construction de paradigmes, l'analyse de Kiparsky (2005) reste floue quant à ce qu'elle affirme réellement en termes d'opération morphologique impliquée dans la passivation en latin. Elle nous laisse avec deux interprétations possibles de l'analyse de Kiparsky (2005). L'une, flexionnelle, permet de générer les formes correctes, mais ne semble pas prendre en compte la spécificité des données indiquant que la passivation ressemble à une opération dérivationnelle en latin. L'autre, dérivationnelle, capte la spécificité des changements sémantiques imprévisibles qui accompagnent la passivation, mais surgénère des formes pour les verbes passifs. Elle n'exprime ainsi pas la spécificité des verbes déponents d'être les seuls à avoir des formes passives en plus de toutes les formes participiales, actives et passives, du gérondif et du supin.

### 1.3 Les « déponents latins », des paradigmes hétéroclites

Pour notre propre analyse, nous nous plaçons dans le cadre d'une approche de type *Mot et Paradigme* et plus précisément inférentielle-réalisationnelle au sens de Stump (2001). Nous nous fixons par ailleurs les objectifs suivants en termes de méthodologie (M) :

- Ma)* Les données particulières permettent d'isoler des questions générales. Dans le cas que nous analysons, cela consiste en particulier à identifier, à la lumière des informations empiriques fournies par la littérature, la nature de l'opération morphologique effectuée par l'opposition actif/passif (en latin) : flexion ou dérivation ?
- Mb)* Une formalisation complète de la génération des formes pour l'intégralité des paradigmes est nécessaire pour garantir la cohérence de l'analyse.
- Mc)* L'implémentation de la formalisation permet de vérifier la génération correcte des formes attendues.
- Md)* Cette implémentation, si elle est faite à grande échelle (avec un lexique à grande échelle associé), permet de valider globalement l'analyse : non seulement les formes des paradigmes analysés au moment de la conception de l'analyse sont correctement générés, mais le modèle couvre aussi de façon satisfaisante tous les paradigmes concernés de la langue étudiée ; il n'y a pas d'oubli de cas particuliers qui autrement pourraient contredire l'analyse. Cela évite également le risque de surévaluer l'importance de phénomènes ponctuels, marginaux par rapport au fonctionnement général de la langue. Dans la mesure où les formes à générer en morphologie flexionnelle sont d'un nombre fini et relativement petit, une telle

VALEUR DU M-PASSIF	ÉTENDUE DU PHÉNOMÈNE
MÉDIO-PASSIF	<i>importante</i>
MUTATIFS	>700 verbes
VERBES DE MOUVEMENT	>500 verbes
IMPERSONNEL	>400 verbes

TABLE 1.14 : Nombre de verbes concernés par les valeurs alternatives du m-passif selon Flobert (1967)

implémentation est tout à fait réalisable et on ne voit donc pas pourquoi on se priverait de la garantie de cohérence qu'elle apporte.

Les données (D) dont notre analyse tente de rendre compte sont les suivantes :

- Da)* Les changements sémantiques opérés par le changement de voix en latin ne sont pas directement prédictibles (cf. section 1.1.2.2).
- Db)* Le s-passif n'est pas réalisé par les seules formes m-passives (cf. section 1.1.2.1). Notamment certaines constructions avec des formes m-actives peuvent exprimer le s-passif.
- Dc)* Les formes des verbes déponents présentent des exposants m-passifs alors que leur fonction est généralement associée au s-actif. Ils sont souvent présentés comme l'exemple type de « *décalage entre forme et fonction* » (Baerman, 2007b).
- Dd)* Les verbes déponents ont un paradigme plus étendu que les autres verbes : ils comportent à la fois des participes actifs et passifs.

Le but de la proposition d'analyse faite dans ce premier chapitre est de montrer la pertinence des exigences méthodologiques formulées ci-dessus. La mise en place d'un système permettant le respect de ces contraintes est présentée dans les chapitres suivants.

### 1.3.1 Le passif latin : dérivation ou flexion ?

Dans la section 1.1.2.2 nous avons exposé l'inventaire des valeurs associées au m-passif tel que l'a présenté Flobert (1967). Parmi ces valeurs on trouve à côté du s-passif attendu, la valeur de médio-passif, de « *mutatif* », de mouvement en cours (dérivé à partir d'un verbe de mouvement initié) et d'impersonnel. Ces valeurs sont chacune associées à un nombre représentatif de verbes. Les nombres (plus ou moins précis) indiqués par Flobert sont récapitulés dans le tableau 1.14.

Rappelons<sup>32</sup> que d'après Wurzel (1989) et Boyé (2011) une telle imprédictibilité est caractéristique des opérations morphologiques de dérivation. L'hypothèse que l'opposition actif/passif puisse relever de la dérivation se justifie également par d'autres arguments que nous détaillons ci-dessous.

Dans les approches lexicalistes de la syntaxe, la structure argumentale d'un verbe fait partie de sa définition dans le lexique : le nombre et le type d'arguments d'un verbe donné sont spécifiés dans le lexique pour chaque lexème (voire, notamment Bresnan (1982); Gazdar et al. (1985); Sag et al. (2003) et Blevins (2003)). La passivation s'analyse alors obligatoirement comme une relation entre deux entrées du lexique. Cette définition du lien entre lexème et structure argumentale a l'avantage de simplifier considérablement l'interface entre la morphologie et la structure argumentale : on a une structure argumentale fixe par lexème qui pourra être correctement implémentée en syntaxe.

En latin, cette approche semble d'autant plus pertinente qu'il semble difficile d'établir un lien direct entre les marques morphologiques portées par une forme et la structure argumentale qu'elle exprime. Loin d'avoir un alignement unique m-actif/s-actif *versus* m-passif/s-passif, les alignements observables sont bien plus complexes, comme le montre notamment Flobert (1967), cf. section 1.1.2.1.

Un autre argument vient par ailleurs directement de la classe de lexèmes m-actifs pouvant exprimer le s-passif. Flobert indique qu'il y a une classe de lexèmes pour lesquels un processus de dérivation, soit à l'aide du suffixe *-esc*, soit par l'intermédiaire d'un changement de la voyelle thématique, résulte en la création d'un nouveau lexème qui, bien que de marquage m-actif, peut exprimer le s-passif<sup>33</sup>. Parmi les exemples cités par Flobert figurent les paires IACEO/IACIO « *reposer* », TENDO/TENEO « *déployer* », PENDO/PENDEO « *peser* », AUGEO/AUGESCO « *renforcer* ». Enfin, les exceptions lexicales observables quant à la défektivité au passif des *activa tantum* ferait de la passivation une opération flexionnelle qui serait au mieux non canonique au sens de Corbett (2007a)

L'opposition actif/passif se comportant comme un processus de dérivation, nous proposons donc l'interprétation suivante des marques morphologiques m-actives/m-passives :

**L'opposition m-actif/m-passif** Le m-actif et le m-passif correspondent à deux classes flexionnelles indépendantes, que nous nommerons désormais respectivement CLASSE A et CLASSE B. Ces deux classes peuvent être reliées par une relation de déri-

---

32. Cité à la section 1.2.1.2.

33. Cf. le point b) dans la liste des marquages du s-passif référencés par Flobert (1967), section 1.1.2.1.

vation. La relation de dérivation en question a en général une fonction de détransitivation et construit des verbes à structure syntaxique s-passive à partir de la base d'un verbe à structure syntaxique s-active, mais cette fonction n'est pas la seule qu'elle puisse assumer.

Un fait supplémentaire qui argumente en faveur de cette approche dérivationnelle de l'opposition m-actif/m-passif est celui mis en évidence par Xu et al. (2007). En effet Xu et al. indiquent que les verbes déponents dénominatifs et déadjectivaux dérivés à partir d'un nom ou adjectif *x* dénotent essentiellement le fait de *se comporter comme un x*, respectivement d'*être x*, tandis que les verbes m-actifs dénominatifs/déadjectivaux ont essentiellement un sens causatif de type *rendre x*. Nous sommes donc en présence de deux règles de dérivation distinctes générant chacune (dans le cas général) une classe de verbes sémantiquement distincts. L'argument essentiel en faveur d'une approche dérivationnelle du passif et de l'existence de classes flexionnelles distinctes pour l'actif et le passif que l'on observe dans ces données est donc le suivant : il n'y a pas que les passifs à être construits par dérivation. Certains verbes déponents sont également générés par des règles dérivationnelles<sup>34</sup>, même si ces règles, construisant des verbes dénominatifs ou déadjectivaux, ont une prédictibilité moindre que la passivation.

Nous représentons dans les tableaux 1.15 et 1.16 pages 69 à 70 une version schématique de deux classes flexionnelles définies pour l'actif et le passif respectivement. Ces tableaux indiquent par ailleurs, par les teintes de rouge plus ou moins intenses, les radicaux employés dans les différents sous-paradigmes construits à partir de chacune des deux classes flexionnelles.

### 1.3.2 Expliquer les paradigmes des verbes déponents

La conséquence d'une telle analyse est que, puisque les formes m-passives sont simplement les formes générées par une seconde classe flexionnelle, il n'y a plus de raison de les considérer comme les formes associées à la valeur s-passive exclusivement. Pour les verbes déponents, il n'y a donc plus lieu de parler de décalage entre forme et fonction. Ce ne sont donc pas des déponents au sens de Baerman (2007b), mais simplement des verbes associés à la classe flexionnelle B. Pour éviter tout amalgame entre la désignation du phénomène théorisé par Baerman (2007b) et le comportement des verbes déponents, nous préférons dans la suite de cette thèse éviter de parler de *déponence* en tant que phénomène non canonique. Nous préférons employer le terme de DÉCALAGE (MORPHOSYNTAXIQUE)

---

34. Construisant des verbes appartenant à la classe flexionnelle B.

CLASSE A						
			IND		SBJV	
			SG	PL	SG	PL
IPFV	PRS	1	RAD1- <i>ō</i>	RAD1- <i>mus</i>	RAD1- <i>am</i>	RAD1- <i>āmus</i>
		2	RAD1- <i>s</i>	RAD1- <i>tis</i>	RAD1- <i>ās</i>	RAD1- <i>ātis</i>
		3	RAD1- <i>t</i>	RAD1- <i>nt</i>	RAD1- <i>at</i>	RAD1- <i>ant</i>
	PST	1	RAD1- <i>bam</i>	RAD1- <i>bāmus</i>	RAD1- <i>rem</i>	RAD1- <i>rēmus</i>
		2	RAD1- <i>bās</i>	RAD1- <i>bātis</i>	RAD1- <i>rēs</i>	RAD1- <i>rētis</i>
		3	RAD1- <i>bat</i>	RAD1- <i>bant</i>	RAD1- <i>ret</i>	RAD1- <i>rent</i>
	FUT	1	RAD1- <i>bo</i>	RAD1- <i>bimus</i>		
		2	RAD1- <i>bis</i>	RAD1- <i>bitis</i>		
		3	RAD1- <i>bit</i>	RAD1- <i>bunt</i>		
PFV	PRS	1	RAD2- <i>i</i>	RAD2- <i>imus</i>	RAD2- <i>erim</i>	RAD2- <i>erimus</i>
		2	RAD2- <i>istī</i>	RAD2- <i>istis</i>	RAD2- <i>eris</i>	RAD2- <i>eritis</i>
		3	RAD2- <i>it</i>	RAD2- <i>ērunt (-ēre)</i>	RAD2- <i>erit</i>	RAD2- <i>ērint</i>
	PST	1	RAD2- <i>eram</i>	RAD2- <i>erāmus</i>	RAD2- <i>issem</i>	RAD2- <i>issēmus</i>
		2	RAD2- <i>erās</i>	RAD2- <i>erātis</i>	RAD2- <i>issēs</i>	RAD2- <i>issētis</i>
		3	RAD2- <i>erat</i>	RAD2- <i>erant</i>	RAD2- <i>isset</i>	RAD2- <i>issent</i>
	FUT	1	RAD2- <i>erō</i>	RAD2- <i>erimus</i>		
		2	RAD2- <i>eris</i>	RAD2- <i>eritis</i>		
		3	RAD2- <i>erit</i>	RAD2- <i>erint</i>		
IMP	FUT	2	RAD1	RAD1- <i>te</i>		
		2	RAD1- <i>to</i>	RAD1- <i>tōte</i>		
		3	RAD1- <i>to</i>	RAD1- <i>tōte</i>		
INF	PRS		RAD1- <i>re</i>			
	PST		RAD2- <i>isse</i>			
	FUT		RAD3- <i>ūrus/-ūra/-ūrum esse</i>			
PTCP	PRS		RAD1- <i>ns, -ntis</i>			
	FUT		RAD3- <i>ūrus/-ūra/-ūrum</i>			
GRV			RAD1- <i>ndī/-ō/-um</i>			
SUP			RAD3- <i>um/-ū</i>			

RAD1

RAD2

RAD3

TABLE 1.15 : Formes de la classe flexionnelle A (verbes actifs canoniques)



CLASSE B						
			IND		SBJV	
			SG	PL	SG	PL
IPFV	PRS	1	RAD1- <i>or</i>	RAD1- <i>mur</i>	RAD1- <i>ar</i>	RAD1- <i>āmur</i>
		2	RAD1- <i>ris (-re)</i>	RAD1- <i>minī</i>	RAD1- <i>āris (-āre)</i>	RAD1- <i>āminī</i>
		3	RAD1- <i>tur</i>	RAD1- <i>ntur</i>	RAD1- <i>atur</i>	RAD1- <i>antur</i>
	PST	1	RAD1- <i>bar</i>	RAD1- <i>bāmur</i>	RAD1- <i>rer</i>	RAD1- <i>rēmur</i>
		2	RAD1- <i>bāris (-bāre)</i>	RAD1- <i>bāminī</i>	RAD1- <i>rēris (-rēre)</i>	RAD1- <i>rēminī</i>
		3	RAD1- <i>bātur</i>	RAD1- <i>bantur</i>	RAD1- <i>rētur</i>	RAD1- <i>rentur</i>
	FUT	1	RAD1- <i>bor</i>	RAD1- <i>bimur</i>		
		2	RAD1- <i>beris (-bere)</i>	RAD1- <i>biminī</i>		
		3	RAD1- <i>bitur</i>	RAD1- <i>buntur</i>		
PFV	PRS	1	RAD3- <i>us sum</i>	RAD3- <i>ī sumus</i>	RAD3- <i>us sim</i>	RAD3- <i>ī sīmus</i>
		2	RAD3- <i>us es</i>	RAD3- <i>ī estis</i>	RAD3- <i>us sīs</i>	RAD3- <i>ī sītis</i>
		3	RAD3- <i>us est</i>	RAD3- <i>ī sunt</i>	RAD3- <i>us sit</i>	RAD3- <i>ī sint</i>
	PST	1	RAD3- <i>us eram</i>	RAD3- <i>ī erāmus</i>	RAD3- <i>us essem</i>	RAD3- <i>ī essēmus</i>
		2	RAD3- <i>us erās</i>	RAD3- <i>ī erātis</i>	RAD3- <i>us essēs</i>	RAD3- <i>ī essētis</i>
		3	RAD3- <i>us erat</i>	RAD3- <i>ī erant</i>	RAD3- <i>us esset</i>	RAD3- <i>ī essent</i>
	FUT	1	RAD3- <i>us erō</i>	RAD3- <i>ī erimus</i>		
		2	RAD3- <i>us eris</i>	RAD3- <i>ī eritis</i>		
		3	RAD3- <i>us erit</i>	RAD3- <i>ī erint</i>		
IMP	PRS	2	RAD1- <i>re</i>	RAD1- <i>minī</i>		
		2	RAD1- <i>tor</i>			
	FUT	3	RAD1- <i>tor</i>	RAD1- <i>ntor</i>		
INF	PRS		RAD1- <i>(r)ī</i>			
	PST		RAD3- <i>us/-a/-um esse</i>			
PTCP	PST		RAD3- <i>us/-a/-um</i>			
GRD			RAD1- <i>ndus/-nda/-ndum</i>			

RAD1

RAD2

RAD3

TABLE 1.16 : Formes de la classe flexionnelle B (verbes passifs canoniques)

dérivé de la définition de Baerman qui définit la déponence comme « *un décalage entre forme et fonction* <sup>35</sup> ». Le terme de DÉPONENT sera gardé uniquement comme une étiquette descriptive réservée à la phénoménologie du latin.

Il reste néanmoins à expliquer la forme inhabituelle des paradigmes des verbes déponents ; comme nous l'avons indiqué au début de ce chapitre, ces verbes possèdent en effet des participes actifs et des gérondifs, appartenant à la classe flexionnelle A tout en ayant par ailleurs des paradigmes formés par la classe flexionnelle B qui comporte des participes passifs. Ce sont, avec les *passiva tantum*, mentionnés notamment par Flobert (1967), les seuls verbes à avoir à la fois des participes présent, futur et passé.

Dans la mesure où les verbes déponents (de même que les *passiva tantum*) possèdent à la fois des formes construites à partir de la classe flexionnelle B et de la classe flexionnelle A, une façon simple de les analyser, cohérente avec une approche dérivationnelle du passif en latin, est de les considérer comme des instances d'*hétéroclise*. Les lexèmes hétéroclites sont des verbes qui construisent leurs paradigmes à partir des formes obtenues par (au moins) deux classes flexionnelles différentes <sup>36</sup>. En l'occurrence, les verbes déponents possèdent la plupart des formes de la classe des verbes passifs, ainsi que les formes non finies des verbes actifs hors infinitifs <sup>37</sup>.

## 1.4 Conception d'une nouvelle analyse formelle de la flexion latine

### 1.4.1 Représentation de l'hétéroclise

Dans cette section nous proposons une nouvelle analyse formelle de la flexion verbale latine. Cette analyse est le fruit de l'application des préceptes méthodologiques définis au début de la section 1.3.

Compte tenu des réponses qui ont pu être données aux questions théoriques à la section 1.3, nous optons pour une approche dérivationnelle du passif en latin. Les tableaux 1.15 et 1.16 pages 69 à 70 que nous présentons ici constituent ainsi deux classes flexionnelles disponibles en latin pour les verbes des première et deuxième conjugaisons.

Dans ces tableaux nous pouvons par ailleurs distinguer les ZONES DE RADICAUX qui constituent un découpage du paradigme d'un verbe en ses sous-paradigmes employant

---

35. « *A mismatch between form and function.* »

36. Voir aussi la définition de l'hétéroclise dans le chapitre 3 et la définition formelle établie au chapitre 4.

37. À savoir les participes présent et futur, le gérondif et le supin.

chacun un des trois radicaux définis pour chaque verbe latin (cf. la section 1.1). Ce que nous appelons ici les « zones de radicaux » correspond aux *partitions* définies par Pirrelli & Battista (2000) ou aux cases de l'*espace thématique* de Bonami & Boyé (2003, 2006). Les zones de radicaux sont représentées par l'emploi d'une couleur spécifique dans les tableaux 1.15 et 1.16 pages 69 à 70.

Ces zones de radicaux sont découpées selon des ESPACES PARTITIONNANTS, c'est-à-dire des ensembles de paquets de traits morphosyntaxiques définissant chacun une des zones de radicaux. Chaque zone de radical définit les règles de formation de radical permettant de former un radical particulier pour l'ensemble des lexèmes utilisant cette zone. Informellement, nous appelons CLASSE DE RADICAUX un regroupement de zones de radicaux partagé par un nombre significatif<sup>38</sup> de lexèmes.

Dans notre analyse du système flexionnel verbal du latin, nous posons ainsi les zones de radicaux découpées par les espaces partitionnants décrits ci-dessous. Elles correspondent aux numéros de radicaux indiqués dans les tableaux 1.15 et 1.16 pages 69 à 70.

Pour les radicaux des verbes associés à la classe A :

**Zone RAD1** La zone RAD1 réalise tous les paquets de traits morphosyntaxiques comportant le trait IPFV pour les formes finies, ainsi que celle de l'infinitif présent (PRS.INF), du participe présent (PRS.PTCP) et du gérondif (GRV).

**Zone RAD2** La zone RAD2 regroupe toutes les structures morphosyntaxiques comportant le trait PFV pour les formes finies et l'infinitif passé (PST.INF).

**Zone RAD3** La zone RAD3 regroupe les structures morphosyntaxiques de l'infinitif et du participe futurs (FUT.INF et FUT.PTCP) et du supin (SUP).

Afin de modéliser la notion d'hétéroclise correspondant notamment à la structure des paradigmes des verbes déponents et des passiva tantum, nous introduisons ici la notion de ZONE FLEXIONNELLE. Les zones flexionnelles sont une généralisation au système des exposants (ou système désinentiel) des zones de radicaux.

Une zone flexionnelle est constituée des règles de réalisation permettant d'engendrer un des sous-paradigmes d'un lexème donné. Plus précisément, une zone flexionnelle est définie par un espace partitionnant associé à un ensemble de règles de réalisation permettant de générer les formes d'un lexème donné pour l'ensemble des paquets de traits morphosyntaxiques de l'espace partitionnant en question. Pour chaque lexème, une

---

38. Pour le système flexionnel étudié. Pour plus d'informations sur la notion de classe flexionnelle dans ce modèle, voir le chapitre 4. Pour plus d'informations sur la notion de découpage cohérent en zones, voir le chapitre 6.

combinaisons spécifique de zones flexionnelles permet de générer l'ensemble des formes de son paradigme.

Une CLASSE FLEXIONNELLE est alors définie à partir des zones flexionnelles comme une combinaison *par défaut* de zones flexionnelles, telle que leur ensemble soit utilisé par un nombre significatif de lexèmes. Les autres combinaisons de zones flexionnelles correspondent à des cas irréguliers (comme l'hétéroclise) à décrire explicitement dans la description du système flexionnel étudié.

Dans notre analyse du système flexionnel verbal du latin, nous stipulons les zones flexionnelles définies par les espaces partitionnants décrits ci-dessous. Les règles de réalisation associées à ces espaces partitionnants dans les zones en question sont représentées par les désinences indiquées dans les tableaux 1.17 et 1.18 pages 74 à 75, pour les zones flexionnelles définissant les classes A et B respectivement. Elles déterminent le découpage différent des paradigmes verbaux selon les propriétés des exposants.

Les zones flexionnelles comportant les désinences de la classe A sont découpées comme suit :

**Zone I-ZONE1** La zone I-ZONE1 regroupe toutes les structures morphosyntaxiques comportant le trait IPFV pour les formes finies.

**Zone I-ZONE2** La zone I-ZONE2 regroupe toutes les structures morphosyntaxiques comportant le trait PFV pour les formes finies.

**Zone I-ZONE3** La zone I-ZONE3 regroupe toutes les structures morphosyntaxiques associées aux formes verbales non finies (ne comportant notamment pas de trait PERSONNE).

Pour les désinences de la classe B, le découpage est le suivant :

**Zone I-ZONE1** La zone I-ZONE1 regroupe toutes les structures morphosyntaxiques comportant le trait IPFV pour les formes finies.

**Zone I-ZONE2** La zone I-ZONE2 regroupe toutes les structures morphosyntaxiques comportant le trait PFV pour les formes finies.

**Zone I-ZONE3** La zone I-ZONE3 regroupe toutes les structures morphosyntaxiques associées aux formes verbales non finies.

On remarque que les découpages des zones de radicaux et des zones de flexion ne se superposent pas nécessairement. Les zones flexionnelles sont illustrées par les couleurs dans les tableaux 1.17 et 1.18 pages 74 à 75.

Les zones flexionnelles permettent de rendre compte de la flexion des différents verbes latins de la façon suivante. Chaque combinaison de zones permet de qualifier la structure

CLASSE A						
			IND		SBJV	
			SG	PL	SG	PL
IPFV	PRS	1	RAD1- <i>ō</i>	RAD1- <i>mus</i>	RAD1- <i>am</i>	RAD1- <i>āmus</i>
		2	RAD1- <i>s</i>	RAD1- <i>tis</i>	RAD1- <i>ās</i>	RAD1- <i>ātis</i>
		3	RAD1- <i>t</i>	RAD1- <i>nt</i>	RAD1- <i>at</i>	RAD1- <i>ant</i>
	PST	1	RAD1- <i>bam</i>	RAD1- <i>bāmus</i>	RAD1- <i>rem</i>	RAD1- <i>rēmus</i>
		2	RAD1- <i>bās</i>	RAD1- <i>bātis</i>	RAD1- <i>rēs</i>	RAD1- <i>rētis</i>
		3	RAD1- <i>bat</i>	RAD1- <i>bant</i>	RAD1- <i>ret</i>	RAD1- <i>rent</i>
	FUT	1	RAD1- <i>bo</i>	RAD1- <i>bimus</i>		
		2	RAD1- <i>bis</i>	RAD1- <i>bitis</i>		
		3	RAD1- <i>bit</i>	RAD1- <i>bunt</i>		
PFV	PRS	1	RAD2- <i>i</i>	RAD2- <i>imus</i>	RAD2- <i>erim</i>	RAD2- <i>erimus</i>
		2	RAD2- <i>istī</i>	RAD2- <i>istis</i>	RAD2- <i>eris</i>	RAD2- <i>eritis</i>
		3	RAD2- <i>it</i>	RAD2- <i>ērunt (-ēre)</i>	RAD2- <i>erit</i>	RAD2- <i>ērint</i>
	PST	1	RAD2- <i>eram</i>	RAD2- <i>erāmus</i>	RAD2- <i>issem</i>	RAD2- <i>issēmus</i>
		2	RAD2- <i>erās</i>	RAD2- <i>erātis</i>	RAD2- <i>issēs</i>	RAD2- <i>issētis</i>
		3	RAD2- <i>erat</i>	RAD2- <i>erant</i>	RAD2- <i>isset</i>	RAD2- <i>issent</i>
	FUT	1	RAD2- <i>erō</i>	RAD2- <i>erimus</i>		
		2	RAD2- <i>eris</i>	RAD2- <i>eritis</i>		
		3	RAD2- <i>erit</i>	RAD2- <i>erint</i>		
IMP	FUT	2	RAD1	RAD1- <i>te</i>		
		2	RAD1- <i>to</i>	RAD1- <i>tōte</i>		
		3	RAD1- <i>to</i>	RAD1- <i>tōte</i>		
INF	PST		RAD1- <i>re</i>			
			RAD2- <i>isse</i>			
			RAD3- <i>ūrus/-ūra/-ūrum esse</i>			
PTCP	PRS		RAD1- <i>ns, -ntis</i>			
	FUT		RAD3- <i>ūrus/-ūra/-ūrum</i>			
GRV	SUP		RAD1- <i>ndī/-ō/-um</i>			
			RAD3- <i>um/-ū</i>			
			I-ZONE1		I-ZONE2	
			I-ZONE3			

TABLE 1.17 : Zones flexionnelles de la classe flexionnelle A

CLASSE B						
			IND		SBJV	
			SG	PL	SG	PL
IPFV	PRS	1	RAD1- <i>or</i>	RAD1- <i>mur</i>	RAD1- <i>ar</i>	RAD1- <i>āmur</i>
		2	RAD1- <i>ris (-re)</i>	RAD1- <i>minī</i>	RAD1- <i>āris (-āre)</i>	RAD1- <i>āminī</i>
		3	RAD1- <i>tur</i>	RAD1- <i>ntur</i>	RAD1- <i>ātur</i>	RAD1- <i>antur</i>
	PST	1	RAD1- <i>bar</i>	RAD1- <i>bāmur</i>	RAD1- <i>rer</i>	RAD1- <i>rēmur</i>
		2	RAD1- <i>bāris (-bāre)</i>	RAD1- <i>bāminī</i>	RAD1- <i>rēris (-rēre)</i>	RAD1- <i>rēminī</i>
		3	RAD1- <i>bātur</i>	RAD1- <i>bantur</i>	RAD1- <i>rētur</i>	RAD1- <i>rentur</i>
	FUT	1	RAD1- <i>bor</i>	RAD1- <i>bimur</i>		
		2	RAD1- <i>beris (-bere)</i>	RAD1- <i>biminī</i>		
		3	RAD1- <i>bitur</i>	RAD1- <i>buntur</i>		
PFV	PRS	1	RAD3- <i>us sum</i>	RAD3- <i>ī sumus</i>	RAD3- <i>us sim</i>	RAD3- <i>ī sīmus</i>
		2	RAD3- <i>us es</i>	RAD3- <i>ī estis</i>	RAD3- <i>us sīs</i>	RAD3- <i>ī sītis</i>
		3	RAD3- <i>us est</i>	RAD3- <i>ī sunt</i>	RAD3- <i>us sit</i>	RAD3- <i>ī sint</i>
	PST	1	RAD3- <i>us eram</i>	RAD3- <i>ī erāmus</i>	RAD3- <i>us essem</i>	RAD3- <i>ī essēmus</i>
		2	RAD3- <i>us erās</i>	RAD3- <i>ī erātis</i>	RAD3- <i>us essēs</i>	RAD3- <i>ī essētis</i>
		3	RAD3- <i>us erat</i>	RAD3- <i>ī erant</i>	RAD3- <i>us esset</i>	RAD3- <i>ī essent</i>
	FUT	1	RAD3- <i>us erō</i>	RAD3- <i>ī erimus</i>		
		2	RAD3- <i>us eris</i>	RAD3- <i>ī eritis</i>		
		3	RAD3- <i>us erit</i>	RAD3- <i>ī erint</i>		
IMP	PRS	2	RAD1- <i>re</i>	RAD1- <i>minī</i>		
	FUT	2	RAD1- <i>tor</i>			
		3	RAD1- <i>tor</i>	RAD1- <i>ntor</i>		
INF	PRS		RAD1- <i>(r)ī</i>			
	PST		RAD3- <i>us/-a/-um esse</i>			
PTCP	PST		RAD3- <i>us/-a/-um</i>			
GRD			RAD1- <i>ndus/-nda/-ndum</i>			
			I-ZONE1		I-ZONE2	
					I-ZONE3	

TABLE 1.18 : Zones flexionnelles de la classe flexionnelle B

d'un type de paradigme donné, appartenant respectivement aux verbes canoniques actifs, aux verbes canoniques passifs, aux verbes déponents et aux DM1 et DM2.

**Verbes de classe A** Les verbes de classe A sont essentiellement les verbes intransitifs et transitifs actifs comme *AMBULO* « *marcher* » ou *AMO* « *aimer* ». Ces verbes utilisent exclusivement des formes construites à partir des zones flexionnelles de la classe A.

**Verbes de classe B** Les verbes de classe B sont essentiellement des verbes passifs dérivés de verbes transitifs actifs de classe A (cf. *AMOR* « *être aimé* », dérivé de *AMO* « *aimer* »). Ces verbes utilisent exclusivement des formes construites à partir des zones flexionnelles de la classe B.

**Verbes hétéroclites de type 1** Les verbes hétéroclites de type 1 comme les verbes déponents, mais aussi les *passiva tantum*, construisent leurs formes à partir de l'intégralité des zones de la classe B auxquelles s'ajoute la zone I-ZONE3 de la classe A.

**Verbes hétéroclites de type 2** Notre modèle nous permet également de rendre compte des verbes déponents mixtes. Les DM1 produisent leurs formes à partir des zones I-ZONE1 et I-ZONE3 de la classe A et des zones I-ZONE2 et I-ZONE3 de la classe B. Parmi ces verbes nous trouvons notamment le verbe *AUGEO* « *augmenter* ». Pour le DM2 *REVERTOR* « *revenir* », les zones utilisées sont les zones I-ZONE2 et I-ZONE3 de la classe A et les zones I-ZONE1 et I-ZONE3 de la classe B.

## 1.4.2 Représentation des niveaux de réalisation des formes

Les classes A et B et les classes de radicaux se superposent par ailleurs à des différenciations fondées sur les variations de la voyelle thématique illustrées dans les tableaux 1.1 et 1.2 à 1.6 ci-dessus. Nous définissons donc, en parallèle aux classes de radicaux et aux classes flexionnelles, les *CLASSES THÉMATIQUES*<sup>39</sup>. Les classes thématiques du latin sont par conséquent les classes associées aux quatre thèmes disponibles : les voyelles A, E et I et les athématiques. Nous les appelons par la suite les classes thématiques TH-A, TH-E, TH-I et TH-C.

Les classes thématiques sont elles aussi, au même titre que les classes de radicaux et les classes flexionnelles, des combinaisons de zones, en l'occurrence de *ZONES THÉMATIQUES*.

---

39. Notons que cette appellation se fait par analogie aux classes de radicaux et aux classes flexionnelles telles que définies plus haut et ne correspond pas directement à la notion de *classe thématique* chez Tribout (2010) où elles dénotent l'ensemble des lexèmes donnant lieu aux mêmes allomorphies radicales. Elles ne sont pas non plus synonymes de la notion de thème utilisée pour caractériser la flexion verbale dans les langues romanes. La notion que nous introduisons ici désigne un niveau d'analyse intermédiaire entre les radicaux et les exposants. Ce niveau est susceptible de pouvoir être employé dans la description de toute langue et ne se limite par conséquent pas à la qualification d'un sous-ensemble de systèmes flexionnels.

Nous les illustrons pour le latin dans les tableaux 1.19 et 1.20 pages 78 à 79. Les zones thématiques sont définies par les deux espaces partitionnants suivants :

**Zone TH-ZONE1** La zone TH-ZONE1<sup>40</sup> regroupe la structure de trait correspondant à 1SG.PRS.IPFV.IND et 3PL.PRS.IPFV.IND ainsi que toutes les structures morphosyntaxiques comportant les combinaisons de traits PRS.IPFV.SBJV, PST.IPFV.IND et FUT.IPFV.IND pour les formes finies ainsi que celle de l'impératif sauf le 1SG.PRS.IMP, du participe présent et du gérondif et de l'adjectif verbal<sup>41</sup>.

**Zone TH-ZONE2** La zone TH-ZONE2<sup>42</sup> regroupe toutes les structures morphosyntaxiques qui ne sont pas dans TH-ZONE1 : l'indicatif présent privé de 1SG.PRS.IPFV.IND et 3PL.PRS.IPFV.IND, le subjonctif imparfait, toutes les formes du perfectif, le 1SG.PRS.IMP, les infinitifs, les participes futurs et passé et le supin.

Dans le cas régulier, un lexème donné n'utilise que des zones d'une seule classe thématique. Dès lors que plusieurs classes sont sollicitées pour l'inventaire de zones pertinent, nous pouvons dire que le paradigme généré est hétéroclite pour le choix de son thème. C'est par exemple le cas de ACCIPIO « *accepter* », comme l'illustre le tableau 1.6 page 39. Les verbes traditionnellement qualifiés de conjugaison « *troisième mixte* » utilisent la zone TH-ZONE1 de la classe thématique TH-I et la zone TH-ZONE2 de la classe TH-C.

Une fois les trois types de zones établis, nous pouvons donc représenter les paradigmes latins comme la combinaison, pour chaque verbe, d'un ensemble de zones de radicaux, d'un ensemble de zones thématiques et d'un ensemble de zones flexionnelles. Les figures 1.13 et 1.14 pages 80 et 81 illustrent la combinaison des classes employées pour la modélisation des paradigmes latins des verbes actifs prototypiques, respectivement pour les verbes passifs prototypiques.

Dans le modèle que nous proposons, un paradigme peut ainsi être hétéroclite en termes d'exposants, de radicaux ou de voyelles thématiques.

À chaque niveau de réalisation de la forme, des phénomènes d'hétéroclise sont susceptibles d'apparaître : pour les verbes de la conjugaison « *troisième mixte* », l'hétéroclise intervient au niveau thème ; pour les verbes déponents, il apparaît au niveau exposant.

---

40. En vert clair dans le tableau 1.19.

41. Selon les verbes, seul le gérondif ou l'adjectif verbal est en général réalisé. Les passiva tantum et les verbes déponents constituent à ce titre des exceptions réalisant les deux formes.

42. En vert foncé dans le tableau 1.20.



CLASSE A						
			IND		SBJV	
			SG	PL	SG	PL
IPFV	PRS	1	RAD1- <i>ō</i>	RAD1- <i>mus</i>	RAD1- <i>am</i>	RAD1- <i>āmus</i>
		2	RAD1- <i>s</i>	RAD1- <i>tis</i>	RAD1- <i>ās</i>	RAD1- <i>ātis</i>
		3	RAD1- <i>t</i>	RAD1- <i>nt</i>	RAD1- <i>at</i>	RAD1- <i>ant</i>
	PST	1	RAD1- <i>bam</i>	RAD1- <i>bāmus</i>	RAD1- <i>rem</i>	RAD1- <i>rēmus</i>
		2	RAD1- <i>bās</i>	RAD1- <i>bātis</i>	RAD1- <i>rēs</i>	RAD1- <i>rētis</i>
		3	RAD1- <i>bat</i>	RAD1- <i>bant</i>	RAD1- <i>ret</i>	RAD1- <i>rent</i>
	FUT	1	RAD1- <i>bo</i>	RAD1- <i>bimus</i>		
		2	RAD1- <i>bis</i>	RAD1- <i>bitis</i>		
		3	RAD1- <i>bit</i>	RAD1- <i>bunt</i>		
PFV	PRS	1	RAD2- <i>i</i>	RAD2- <i>imus</i>	RAD2- <i>erim</i>	RAD2- <i>erimus</i>
		2	RAD2- <i>istī</i>	RAD2- <i>istis</i>	RAD2- <i>eris</i>	RAD2- <i>eritis</i>
		3	RAD2- <i>it</i>	RAD2- <i>ērunt (-ēre)</i>	RAD2- <i>erit</i>	RAD2- <i>ērint</i>
	PST	1	RAD2- <i>eram</i>	RAD2- <i>erāmus</i>	RAD2- <i>issem</i>	RAD2- <i>issēmus</i>
		2	RAD2- <i>erās</i>	RAD2- <i>erātis</i>	RAD2- <i>issēs</i>	RAD2- <i>issētis</i>
		3	RAD2- <i>erat</i>	RAD2- <i>erant</i>	RAD2- <i>isset</i>	RAD2- <i>issent</i>
	FUT	1	RAD2- <i>erō</i>	RAD2- <i>erimus</i>		
		2	RAD2- <i>eris</i>	RAD2- <i>eritis</i>		
		3	RAD2- <i>erit</i>	RAD2- <i>erint</i>		
IMP	PRS FUT	2	RAD1	RAD1- <i>te</i>		
		2	RAD1- <i>to</i>	RAD1- <i>tōte</i>		
		3	RAD1- <i>to</i>	RAD1- <i>tōte</i>		
INF	PRS PST FUT		RAD1- <i>re</i>			
			RAD2- <i>isse</i>			
			RAD3- <i>ūrus/-ūra/-ūrum esse</i>			
PTCP	PRS FUT		RAD1- <i>ns, -ntis</i>			
			RAD3- <i>ūrus/-ūra/-ūrum</i>			
GRV SUP			RAD1- <i>ndī/-ō/-um</i>			
			RAD3- <i>um/-ū</i>			

TH-ZONE1

TH-ZONE2

TABLE 1.19 : Zones thématiques de la classe flexionnelle A

CLASSE B						
			IND		SBJV	
			SG	PL	SG	PL
IPFV	PRS	1	RAD1- <i>or</i>	RAD1- <i>mur</i>	RAD1- <i>ar</i>	RAD1- <i>āmur</i>
		2	RAD1- <i>ris (-re)</i>	RAD1- <i>minī</i>	RAD1- <i>āris (-āre)</i>	RAD1- <i>āminī</i>
		3	RAD1- <i>tur</i>	RAD1- <i>ntur</i>	RAD1- <i>ātur</i>	RAD1- <i>antur</i>
	PST	1	RAD1- <i>bar</i>	RAD1- <i>bāmur</i>	RAD1- <i>rer</i>	RAD1- <i>rēmur</i>
		2	RAD1- <i>bāris (-bāre)</i>	RAD1- <i>bāminī</i>	RAD1- <i>rēris (-rēre)</i>	RAD1- <i>rēminī</i>
		3	RAD1- <i>bātur</i>	RAD1- <i>bantur</i>	RAD1- <i>rētur</i>	RAD1- <i>rentur</i>
	FUT	1	RAD1- <i>bor</i>	RAD1- <i>bimur</i>		
		2	RAD1- <i>beris (-bere)</i>	RAD1- <i>biminī</i>		
		3	RAD1- <i>bitur</i>	RAD1- <i>buntur</i>		
PFV	PRS	1	RAD3- <i>us sum</i>	RAD3- <i>ī sumus</i>	RAD3- <i>us sim</i>	RAD3- <i>ī sīmus</i>
		2	RAD3- <i>us es</i>	RAD3- <i>ī estis</i>	RAD3- <i>us sīs</i>	RAD3- <i>ī sītis</i>
		3	RAD3- <i>us est</i>	RAD3- <i>ī sunt</i>	RAD3- <i>us sit</i>	RAD3- <i>ī sint</i>
	PST	1	RAD3- <i>us eram</i>	RAD3- <i>ī erāmus</i>	RAD3- <i>us essem</i>	RAD3- <i>ī essēmus</i>
		2	RAD3- <i>us erās</i>	RAD3- <i>ī erātis</i>	RAD3- <i>us essēs</i>	RAD3- <i>ī essētis</i>
		3	RAD3- <i>us erat</i>	RAD3- <i>ī erant</i>	RAD3- <i>us esset</i>	RAD3- <i>ī essent</i>
	FUT	1	RAD3- <i>us erō</i>	RAD3- <i>ī erimus</i>		
		2	RAD3- <i>us eris</i>	RAD3- <i>ī eritis</i>		
		3	RAD3- <i>us erit</i>	RAD3- <i>ī erint</i>		
IMP	PRS	2	RAD1- <i>re</i>	RAD1- <i>minī</i>		
		2	RAD1- <i>tor</i>			
	FUT	3	RAD1- <i>tor</i>	RAD1- <i>ntor</i>		
INF	PRS		RAD1-( <i>r</i> ) <i>ī</i>			
	PST		RAD3- <i>us/-a/-um esse</i>			
PTCP	PST		RAD3- <i>us/-a/-um</i>			
GRD			RAD1- <i>ndus/-nda/-ndum</i>			
			TH-ZONE1	TH-ZONE2		

TABLE 1.20 : Zones thématiques de la classe flexionnelle B

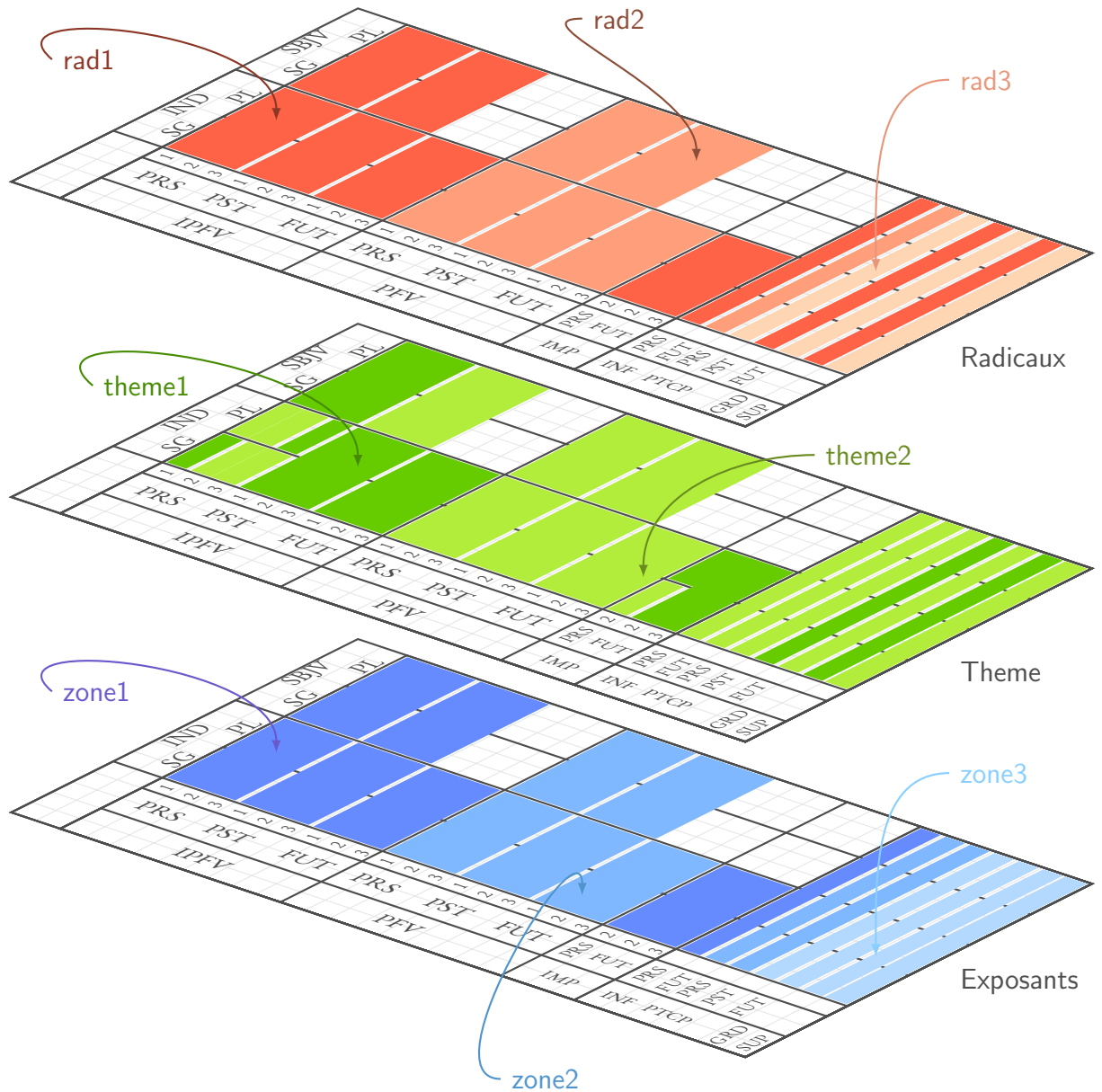


FIGURE 1.13 : Combinaison standard de zones pour les verbes actifs prototypiques

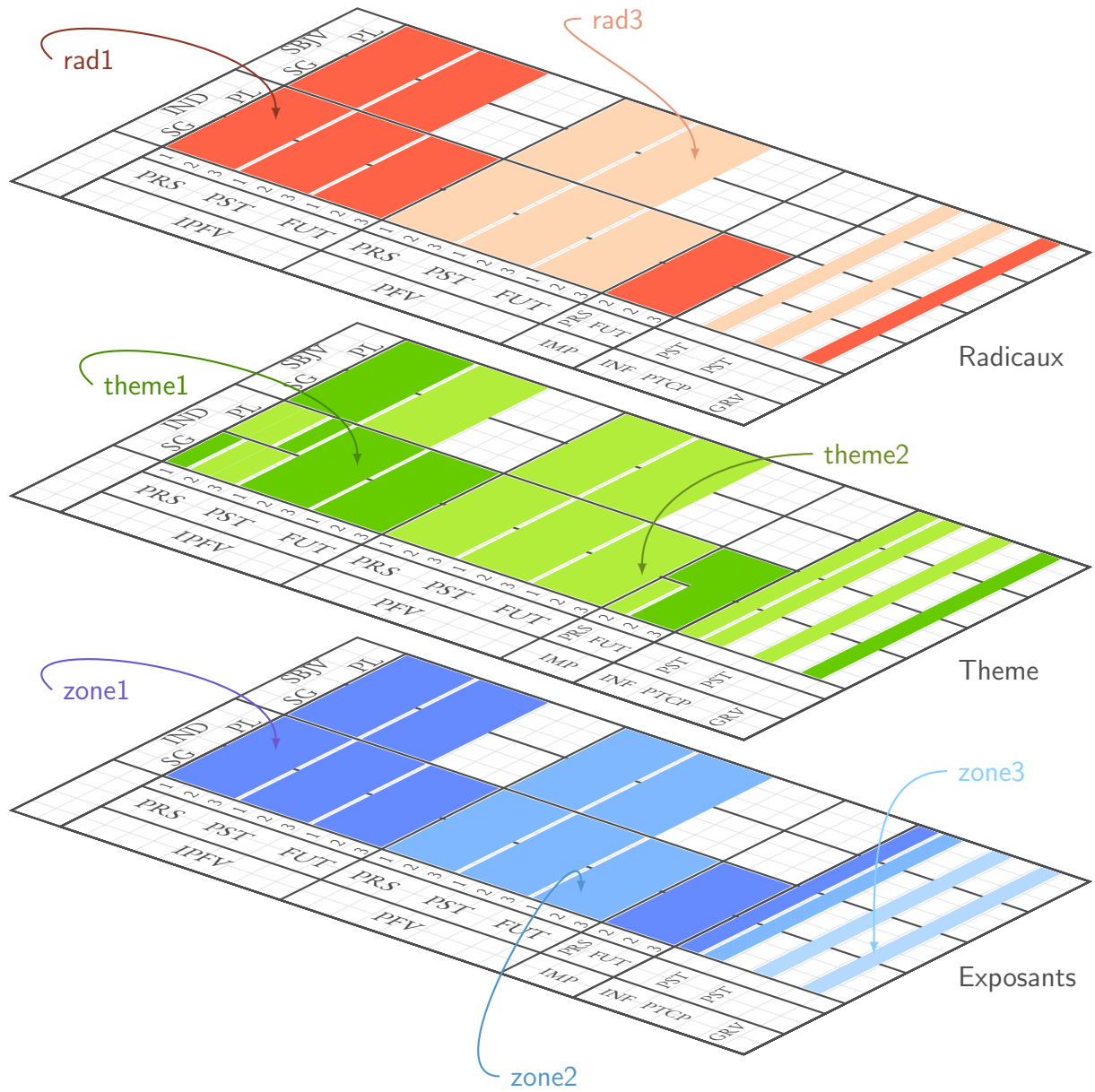


FIGURE 1.14 : Combinaison standard de zones pour les verbes passifs prototypiques

### 1.4.3 Classes flexionnelles des verbes athématiques et à thème en *i*

Les classes A et B ne sont néanmoins pas suffisantes pour représenter l'ensemble des paradigmes verbaux latins. En effet, les verbes appartenant aux classes de la troisième et de la quatrième conjugaison (y compris la *troisième mixte*) suivent un inventaire de désinences légèrement différent de celui indiqué dans les tableaux 1.17 et 1.18 pages 74 et 75. Il nous faut donc poser deux classes flexionnelles supplémentaires au niveau exposant, à savoir les classes C et D, illustrées par les tableaux 1.21 et 1.22 <sup>43</sup> pages 83 et 84. Notons cependant que quand bien même les classes A et C, respectivement B et D, diffèrent formellement dans l'inventaire de leurs désinences, leur découpage en zones flexionnelles est néanmoins soumis aux mêmes espaces partitionnants définis plus haut. Par ailleurs, les verbes appartenant aux classes C et D ne diffèrent des verbes appartenant aux classes A et B qu'au niveau exposant. Les verbes appartenant aux classes A et C, respectivement B et D, sont identiques aux niveaux radical et thématique.

### 1.4.4 Modélisation des règles de réalisation

Nous avons implémenté l'analyse décrite dans les sections 1.4.1 à 1.4.3 dans le formalisme Alexina (Sagot, 2010). Nous avons également associé notre analyse au lexique *Leffla* développé pour l'occasion <sup>44</sup> et dont les entrées ont été extraites des 1 475 verbes extraits du lexique de Prior & Wohlberg (1995) ainsi que de données complémentaires du Wiktionary <sup>45</sup> :

Notre lexique *Leffla* complet comporte ainsi les entrées lexicales non défectives obtenues de la façon suivante :

- Les 1 003 entrées transitives complètes de Prior & Wohlberg (1995) ont permis de construire 1 003 entrées appartenant aux classes A et C d'une part et 1 003 entrées dérivées associées respectivement aux classes B et D de l'autre.
- Les 109 verbes défectifs au seul passif chez Prior & Wohlberg (1995) sont autant de verbes complets des classes A et C dans *Leffla* et correspondent par ailleurs à 109 verbes défectifs de classes B et D.
- De même les 22 autres verbes défectifs intransitifs de Prior & Wohlberg (1995) se retrouvent dans les classes A et C, chacun de ces verbes ne réalisant pas l'intégralité

---

43. Les désinences qui diffèrent sont notées en rouge dans les tableaux.

44. Notre lexique du latin *Leffla* a été développé dans une version adaptée du formalisme Alexina décrite plus en détail dans le chapitre 5.

45. Dictionnaire collaboratif en ligne, semi-structuré, disponible à l'adresse <http://en.wiktionary.org> (valable au 11 juin 2013).

CLASSE C						
			IND		SBJV	
			SG	PL	SG	PL
IPFV	PRS	1	RAD1- <i>ō</i>	RAD1- <i>mus</i>	RAD1- <i>am</i>	RAD1- <i>āmus</i>
		2	RAD1- <i>s</i>	RAD1- <i>tis</i>	RAD1- <i>ās</i>	RAD1- <i>ātis</i>
		3	RAD1- <i>t</i>	RAD1- <i>unt</i>	RAD1- <i>at</i>	RAD1- <i>ant</i>
	PST	1	RAD1- <i>ēbam</i>	RAD1- <i>ēbāmus</i>	RAD1- <i>rem</i>	RAD1- <i>rēmus</i>
		2	RAD1- <i>ebās</i>	RAD1- <i>ebātis</i>	RAD1- <i>rēs</i>	RAD1- <i>rētis</i>
		3	RAD1- <i>ēbat</i>	RAD1- <i>ēbant</i>	RAD1- <i>ret</i>	RAD1- <i>rent</i>
	FUT	1	RAD1- <i>am</i>	RAD1- <i>ēmus</i>		
		2	RAD1- <i>ēs</i>	RAD1- <i>ētis</i>		
		3	RAD1- <i>et</i>	RAD1- <i>ent</i>		
PFV	PRS	1	RAD2- <i>i</i>	RAD2- <i>imus</i>	RAD2- <i>erim</i>	RAD2- <i>erimus</i>
		2	RAD2- <i>istī</i>	RAD2- <i>istis</i>	RAD2- <i>eris</i>	RAD2- <i>eritis</i>
		3	RAD2- <i>it</i>	RAD2- <i>ērunt (-ēre)</i>	RAD2- <i>erit</i>	RAD2- <i>erint</i>
	PST	1	RAD2- <i>eram</i>	RAD2- <i>erāmus</i>	RAD2- <i>issem</i>	RAD2- <i>issēmus</i>
		2	RAD2- <i>erās</i>	RAD2- <i>erātis</i>	RAD2- <i>issēs</i>	RAD2- <i>issētis</i>
		3	RAD2- <i>erat</i>	RAD2- <i>erant</i>	RAD2- <i>isset</i>	RAD2- <i>issent</i>
	FUT	1	RAD2- <i>erō</i>	RAD2- <i>erimus</i>		
		2	RAD2- <i>eris</i>	RAD2- <i>eritis</i>		
		3	RAD2- <i>erit</i>	RAD2- <i>erint</i>		
IMP	FUT	2	RAD1	RAD1- <i>te</i>		
		2	RAD1- <i>to</i>	RAD1- <i>tōte</i>		
		3	RAD1- <i>to</i>	RAD1- <i>tōte</i>		
INF	PST		RAD1- <i>re</i>			
			RAD2- <i>isse</i>			
			RAD3- <i>ūrus/-ūra/-ūrum esse</i>			
PTCP	PRS		RAD1- <i>ns, -ntis</i>			
			RAD3- <i>ūrus/-ūra/-ūrum</i>			
GRV	FUT		RAD1- <i>ndī/-ō/-um</i>			
			RAD3- <i>um/-ū</i>			
SUP						

I-ZONE1

I-ZONE2

I-ZONE3

TABLE 1.21 : Zones flexionnelles de la classe flexionnelle c

CLASSE D						
			IND		SBJV	
			SG	PL	SG	PL
IPFV	PRS	1	RAD1- <i>or</i>	RAD1- <i>mur</i>	RAD1- <i>ar</i>	RAD1- <i>āmur</i>
		2	RAD1- <i>ris (-re)</i>	RAD1- <i>minī</i>	RAD1- <i>āris (-āre)</i>	RAD1- <i>āminī</i>
		3	RAD1- <i>tur</i>	RAD1- <i>untur</i>	RAD1- <i>ātur</i>	RAD1- <i>antur</i>
	PST	1	RAD1- <i>ēbar</i>	RAD1- <i>ēbāmur</i>	RAD1- <i>rer</i>	RAD1- <i>rēmur</i>
		2	RAD1- <i>ēbāris (-ēbāre)</i>	RAD1- <i>ēbāminī</i>	RAD1- <i>rēris (-rēre)</i>	RAD1- <i>rēminī</i>
		3	RAD1- <i>ēbātur</i>	RAD1- <i>ēbantur</i>	RAD1- <i>rētur</i>	RAD1- <i>rentur</i>
	FUT	1	RAD1- <i>ar</i>	RAD1- <i>ēmur</i>		
		2	RAD1- <i>ēris (-ēre)</i>	RAD1- <i>ēminī</i>		
		3	RAD1- <i>ētur</i>	RAD1- <i>entur</i>		
PFV	PRS	1	RAD3- <i>us sum</i>	RAD3- <i>ī sumus</i>	RAD3- <i>us sim</i>	RAD3- <i>ī sīmus</i>
		2	RAD3- <i>us es</i>	RAD3- <i>ī estis</i>	RAD3- <i>us sīs</i>	RAD3- <i>ī sītis</i>
		3	RAD3- <i>us est</i>	RAD3- <i>ī sunt</i>	RAD3- <i>us sit</i>	RAD3- <i>ī sint</i>
	PST	1	RAD3- <i>us eram</i>	RAD3- <i>ī erāmus</i>	RAD3- <i>us essem</i>	RAD3- <i>ī essēmus</i>
		2	RAD3- <i>us erās</i>	RAD3- <i>ī erātis</i>	RAD3- <i>us essēs</i>	RAD3- <i>ī essētis</i>
		3	RAD3- <i>us erat</i>	RAD3- <i>ī erant</i>	RAD3- <i>us esset</i>	RAD3- <i>ī essent</i>
	FUT	1	RAD3- <i>us erō</i>	RAD3- <i>ī erimus</i>		
		2	RAD3- <i>us eris</i>	RAD3- <i>ī eritis</i>		
		3	RAD3- <i>us erit</i>	RAD3- <i>ī erint</i>		
IMP	PRS FUT	2	RAD1- <i>re</i>	RAD1- <i>minī</i>		
		2	RAD1- <i>tor</i>			
		3	RAD1- <i>tor</i>	RAD1- <i>ntor</i>		
INF	PRS PST		RAD1-( <i>r</i> ) <i>ī</i>			
			RAD3- <i>us/-a/-um esse</i>			
		PTCP GRD	PST		RAD3- <i>us/-a/-um</i>	
	RAD1- <i>ndus/-nda/-ndum</i>					
			I-ZONE1		I-ZONE2	
					I-ZONE3	

TABLE 1.22 : Zones flexionnelles de la classe flexionnelle D

des cases définies par la classe à laquelle il appartient.

- Les 193 entrées intransitives de Prior & Wohlberg (1995) correspondent à autant d'entrées appartenant aux classes A et C dans *Leffa*.
- Les 137 verbes déponents de Prior & Wohlberg (1995) sont dans *Leffa* des entrées hétéroclites appartenant majoritairement aux classes B et D et empruntant leurs formes supplémentaires aux classes A et C respectivement.
- De même, les 7 verbes déponents mixtes sont des verbes déponents construits à partir des classes A et B et des classes C et D.
- Enfin, les 4 déponents défectifs ne figurent pas dans *Leffa*, mais ils constituent des entrées hétéroclites appartenant majoritairement aux classes B et D et empruntant leurs formes supplémentaires aux classes A et C respectivement sans cependant réaliser l'intégralité des formes.

Après cette conversion, notre lexique comportait 2 587 entrées. Nous avons filtré ces entrées à l'aide du Wiktionary pour aboutir finalement à un lexique du latin comportant 2 271 entrées. Les différents types de verbes sont résumés au tableau 1.23.

TYPE	NOMBRE	POURCENTAGE
TOTAL	2 271	100,0%
CLASSE A/C	1 065	46,9%
dont complets	1 010	44,5%
dont défectifs	55	2,4%
CLASSE B/D	976	43,0%
dont complets	968	42,6%
dont défectifs	8	0,4%
HÉTÉROCLITES	166	7,3%
dont VERBES DÉPONENTS <sup>46</sup>	159	7,0%
dont VERBES DÉPONENTS MIXTES	7	0,3%
dont défectifs	1	0,04%

TABLE 1.23 : Répartition des types de verbes dans *Leffa*

Par ailleurs, notre implémentation ne s'intéresse ici qu'aux règles de formation des formes synthétiques. Les formes périphrastiques sont considérées comme relevant de procédés distincts que nous ne détaillons pas plus à ce stade.

46. Et passiva tantum.



Nous rendons compte de notre implémentation par les tableaux 1.25 à 1.27 pages 88 à 89 qui simulent un inventaire de classes de position (PFM, Stump (2001))<sup>47</sup>. Les règles porte-manteau qui réalisent deux positions à la fois sont représentées dans une case unique couvrant les deux positions d'exponence disponibles indiquées dans les tableaux 1.25 à 1.28 pages 88 à 89. Ces derniers indiquent pour chaque exposant quelles sont les valeurs de traits pour lesquelles il doit être réalisé sur une forme donnée. Dans les cas où plusieurs paquets de traits morphosyntaxiques peuvent être réalisés par la même règle, ces paquets de traits morphosyntaxiques sont indiqués à la suite, séparés par « | ». Les tableaux de règles de réalisation correspondent respectivement aux classes flexionnelles A, B, C et D définies plus haut. Les différences entre A et C d'une part et B et D de l'autre sont indiquées en rouge dans les tableaux.

Les traits retenus pour l'analyse sont résumés dans le tableau 1.24 page 87.

---

47. Les règles PFM correspondantes sont données en annexe A.1. Le formalisme *Paradigm Function Morphology* (Stump, 2001) est par ailleurs décrit de façon plus circonstanciée au chapitre 2.

	PERSONNE	NOMBRE	TEMPS	ASPECT	MODE
INDICATIF PRÉSENT	1 ∨ 2 ∨ 3	sg ∨ pl	prs	ipfv	ind
SUBJONCTIF PRÉSENT	1 ∨ 2 ∨ 3	sg ∨ pl	prs	ipfv	sbjv
IMPÉRATIF PRÉSENT	2	sg ∨ pl	prs	ipfv	imp
INFINITIF PRÉSENT			prs	ipfv	inf
PARTICIPE PRÉSENT	case	sg ∨ pl	prs	ipfv	part
INDICATIF IMPARFAIT	1 ∨ 2 ∨ 3	sg ∨ pl	pst	ipfv	ind
SUBJONCTIF IMPARFAIT	1 ∨ 2 ∨ 3	sg ∨ pl	pst	ipfv	sbjv
INDICATIF FUTUR	1 ∨ 2 ∨ 3	sg ∨ pl	fut	ipfv	ind
IMPÉRATIF FUTUR	2 ∨ 3	sg ∨ pl	fut	ipfv	imp
INFINITIF FUTUR			fut	ipfv	inf
PARTICIPE FUTUR	(+cas)	sg ∨ pl	fut	ipfv	part
INDICATIF PARFAIT	1 ∨ 2 ∨ 3	sg ∨ pl	prs	pfv	ind
SUBJONCTIF PARFAIT	1 ∨ 2 ∨ 3	sg ∨ pl	prs	pfv	sbjv
IND. PLUS-QUE-PARFAIT	1 ∨ 2 ∨ 3	sg ∨ pl	pst	pfv	ind
SUBJ. PLUS-QUE-PARFAIT	1 ∨ 2 ∨ 3	sg ∨ pl	pst	pfv	sbjv
PASSÉ ANTÉRIEUR	1 ∨ 2 ∨ 3	sg ∨ pl	fut	pfv	ind
INFINITIF PASSÉ			pst	pfv	inf
PARTICIPE PASSÉ	(+cas, genre)	sg ∨ pl	pst	pfv	ptcp
GÉRONDIF	(+cas, genre)				grv
ADJ. VERBAL	(+cas, genre)	sg ∨ pl			grd
SUPIN	(+cas, genre)				sup

TABLE 1.24 : Inventaire des traits morphosyntaxiques intervenant dans l'analyse

RADICAL	1	2
RAD1 RAD2 RAD3	PRS.IPFV.INF : <i>re</i> 2.SG.PRS.IPFV.IMP : $\emptyset$ 3PL.FUT.IPFV.IMP : <i>nto</i> PST.PFV.INF : <i>isse</i> PRS.PFV.IND : <i>i</i> 2.SG.PRS.PFV.IND : <i>istī</i> 2PL.PRS.PFV.IND : <i>istīs</i> NOM.SG.PRS.PTCP : <i>ns</i>	
	PRS.IPFV.SBJV : <i>a</i> PST.IPFV.IND : <i>ba</i> PST.IPFV.SBJV : <i>re</i> FUT.IPFV.IND : <i>bi</i> FUT.IPFV.IMP : <i>to</i> PRS.PFV.IND : <i>i</i> PST.PFV.IND : <i>era</i> PST.PFV.SBJV : <i>isse</i> AUTRES : $\emptyset$ FUT.PFV.IND PRS.PFV.SBJV : <i>eri</i> PRS.PTCP : <i>nt</i> FUT.PTCP : <i>ur</i> GRV : <i>nd</i>	1SG.PRS.IND 1SG.FUT.IND : <i>o</i> 1SG.SBJV 1SG.PST.IND : <i>m</i> 2.SG : <i>s</i> 3SG : <i>t</i> 1PL : <i>mus</i> 2PL : <i>tis</i> 3PL : <i>nt</i> 2PL.IPFV.IMP : <i>te</i> Cas : <i>us/a/um/...</i> Cas : <i>Ø/em/is/...</i>

TABLE 1.25 : Classes de position pour les verbes de type A

RADICAL	1	2
RAD1 RAD3	PRS.IPFV.INF : <i>ri</i> 2.SG.PRS.IPFV.IMP : <i>re</i> 2PL.PRS.IPFV.IMP : <i>minī</i> 2.SG.FUT.IPFV.IMP 3SG.FUT.IPFV.IMP : <i>tor</i>	
	PRS.IPFV.SBJV : <i>a</i> PST.IPFV.IND : <i>ba</i> PST.IPFV.SBJV : <i>re</i> FUT.IPFV.IND : <i>be</i> GRD : <i>nd</i> AUTRES : $\emptyset$	1SG.IPFV.IND : <i>or</i> 1.IPFV.SG : <i>r</i> 2.IPFV.SG : <i>ris</i> 3.IPFV.SG : <i>tur</i> 1.IPFV.PL : <i>mur</i> 2.IPFV.PL : <i>minī</i> 3.IPFV.PL : <i>ntur</i> Cas : <i>us/a/um/...</i>

TABLE 1.26 : Classes de position pour les verbes de type B

RADICAL	1	2
RAD1 RAD2 RAD3	PRS.IPFV.INF : <i>re</i> 2.SG.PRS.IPFV.IMP : $\emptyset$ 3PL.FUT.IPFV.IMP : <i>nto</i> PST.PFV.INF : <i>isse</i> PRS.PFV.IND : <i>i</i> 2.SG.PRS.PFV.IND : <i>istī</i> 2PL.PRS.PFV.IND : <i>istīs</i> NOM.SG.PRS.PTCP : <i>ens</i>	
	PRS.IPFV.SBJV : <i>a</i> PST.IPFV.IND : <i>eba</i> PST.IPFV.SBJV : <i>re</i> FUT.IPFV.IND : <i>e</i> FUT.IPFV.IMP : <i>to</i> PRS.PFV.IND : <i>i</i> PST.PFV.IND : <i>era</i> PST.PFV.SBJV : <i>isse</i> AUTRES : $\emptyset$ FUT.PFV.IND PRS.PFV.SBJV : <i>eri</i> PRS.PTCP : <i>ent</i> FUT.PTCP : <i>ur</i> GRV : <i>end</i>	1SG.PRS.IND 1SG.FUT.IND : <i>o</i> 1SG.SBJV 1SG.PST.IND : <i>m</i> 2.SG : <i>s</i> 3SG : <i>t</i> 1PL : <i>mus</i> 2PL : <i>tis</i> 3PL : <i>nt</i> 2PL.IPFV.IMP : <i>te</i> Cas : <i>us/a/um/...</i> Cas : <i>Ø/em/is/...</i>

TABLE 1.27 : Classes de position pour les verbes de type c

RADICAL	1	2
RAD1 RAD3	PRS.IPFV.INF : <i>ri</i> 2.SG.PRS.IPFV.IMP : <i>re</i> 2PL.PRS.IPFV.IMP : <i>minī</i> 2.SG.FUT.IPFV.IMP 3SG.FUT.IPFV.IMP : <i>tor</i>	
	PRS.IPFV.SBJV 1SG.FUT.IPFV.IND : <i>a</i> PST.IPFV.IND : <i>eba</i> PST.IPFV.SBJV : <i>re</i> FUT.IPFV.IND : <i>e</i> GRD : <i>end</i> AUTRES : $\emptyset$	1SG.IPFV.IND : <i>or</i> 1.IPFV.SG : <i>r</i> 2.IPFV.SG : <i>ris</i> 3.IPFV.SG : <i>tur</i> 1.IPFV.PL : <i>mur</i> 2.IPFV.PL : <i>minī</i> 3.IPFV.PL : <i>ntur</i> Cas : <i>us/a/um/...</i>

TABLE 1.28 : Classes de position pour les verbes de type d

### 1.4.5 La défektivité des verbes latins

Au delà des problèmes traités jusqu'ici, le latin affiche par ailleurs de nombreux cas de défektivité. En effet, il y a un grand nombre de verbes pour lesquels aucune réalisation n'est associée à certains paquets de traits morphosyntaxiques, pourtant habituellement exprimables par les entrées lexicales verbales. Dans *Leffa*, la défektivité concerne 351 verbes, c'est-à-dire presque un dixième du lexique.

Un cas de défektivité extrême concerne le verbe *INQUAM* « dire ». Ce verbe possède en tout douze formes, à savoir le sous-paradigme complet de l'indicatif présent imperfectif, la troisième personne du singulier de l'imparfait (PST.IPFV), les deuxième et troisième personnes du singulier du futur imperfectif et la première et la deuxième personnes du parfait (PRS.PFV). Ce type de défektivité est propre à *INQUAM*. Elle ne se retrouve pour aucune autre entrée lexicale verbale.

Mais il existe également des cas de défektivité qui concernent un plus grand nombre de lexèmes. Ainsi, parmi les 2 587 verbes obtenus à partir du lexique de Prior & Wohlberg (1995), 85 verbes (3,3%) ne sont définis que pour les troisièmes personnes du singulier, les infinitifs (et parfois les participes). Deux exemples de ce type de verbes sont le verbe *LICET* « être permis » de la classe A et le verbe *ACCEDITUR* « approcher » de la classe D.

Afin de rendre compte des diverses structures des paradigmes latins, nous devons donc également pouvoir rendre compte de la défektivité qui affecte nombre de paradigmes. Les inventaires, parfois arbitraires, de structures morphosyntaxiques réalisables doivent pouvoir être définis pour les verbes qui ne réalisent pas par défaut l'intégralité des traits tels que définis dans le tableau 1.24 page 87. Cette définition peut se faire directement pour chaque entrée ou par l'appartenance d'un lexème à un type de défektivité défini dans la grammaire.

Pour les deux exemples ci-dessus, nous proposons donc d'un côté de définir dans la description morphologique (dans la grammaire) un groupe de verbes dits *verbes impersonnels* qui ne réalisent que les troisièmes personnes du singulier, les infinitifs (et parfois les participes). Les verbes comme *LICET* et *ACCEDITUR* appartiennent alors au groupe des verbes impersonnels. De l'autre côté, nous proposons de stipuler les traits réellement réalisables dans le lexique pour les verbes hautement irréguliers comme *INQUAM*. Cela évite de multiplier les définitions de types de défektivité dans la grammaire et d'alourdir inutilement la description.

En plus de définitions de notions comme les classes flexionnelles, les classes de radicaux et les classes thématiques, nous avons ainsi besoin d'une définition explicite de ce qui constitue une entrée lexicale et de la façon dont peuvent être stipulées dans le lexique les

irrégularités dans les paradigmes qui ne touchent que très peu d'unités lexicales.

Ainsi, nous pouvons représenter chaque entrée lexicale par sa forme de citation, les différentes zones qui sont employées pour construire son paradigme et, dans les cas où un lexème donné ne réalise pas l'intégralité des structures morphosyntaxiques attendues compte tenu de son appartenance à une catégorie morphosyntaxique donnée, l'inventaire des structures morphosyntaxiques qu'il réalise (soit par appartenance à un groupe particulier, soit directement dans le lexique). Le tableau 1.29 page 92 montre pour six verbes latins, un résumé des informations à spécifier pour chaque entrée lexicale<sup>48</sup>. *Standard* dans la colonne TRAITs RÉALISÉS correspond à la réalisation de l'intégralité des combinaisons de traits définies dans les tableau 1.24 page 87. *Impersonnel* correspond à la réalisation de toutes les combinaisons de traits du tableau 1.24 comportant également les traits 3SG et INF. Cette définition de *impersonnel* peut être faite dans la grammaire.

#### 1.4.6 Extension formelle de l'analyse

Une critique possible de notre analyse<sup>49</sup> est qu'elle semble comporter une redondance non négligeable dans la définition des classes flexionnelles A et B d'une part et C et D de l'autre : chaque instance de la dérivation passive<sup>50</sup> correspond à un changement de classe flexionnelle vers la classe B pour les verbes de classe A et vers la classe D pour les verbes de classe C. Or cette systématisme est déjà captée par la nature thématique des verbes concernés. On aimerait donc pouvoir simplifier le système de classes flexionnelles et mieux rendre compte de la correspondance immédiate entre thèmes en A et en E et les classes A et B d'une part et les thèmes en I, athématique et mixte et les classes C et D de l'autre.

Grâce au modèle que nous avons esquissé ici et que nous introduisons formellement au chapitre 4, nous avons la possibilité de réduire les quatre classes flexionnelles A, B, C et D que nous avons stipulées pour rendre compte de la flexion verbale latine à seulement deux classes. Le modèle permet en effet de subdiviser chacun des niveaux (radical, thème, exponence) en plusieurs niveaux de ce type.

Ainsi, si l'on s'autorise l'introduction d'un niveau exposant supplémentaire, nous pourrions stipuler un niveau exposant correspondant au bloc 1 de l'analyse que nous avons donnée dans ce chapitre et un autre niveau qui correspondrait au bloc 2. Nous aurions alors

---

48. La dénomination des zones suit celles données dans les tableaux ci-dessus. Par exemple la zone A-I-ZONE1 correspond à la zone I-ZONE1 dans la classe flexionnelle A.

49. Nous remercions Grev Corbett et Matthew Baerman pour nous avoir fait cette suggestion lors d'une présentation au sein d'un atelier du SEMM.

50. Sauf pour les exemples de passivation par changement de thème indiqués par Flobert (1967) et mentionnés plus haut.

FORME DE CITATION	CATÉGORIE	ZONES DE RADICAUX	ZONES THÉMATIQUES	ZONES FLEXIONNELLES	TRAITTS RÉALISÉS
AMO	VERBE	RAD1	A-TH-ZONE1	A-1-ZONE1	<i>standard</i>
		RAD2	A-TH-ZONE2	A-1-ZONE2	
		RAD3		A-1-ZONE3	
ACCIPIO	VERBE	RAD1	I-TH-ZONE1	C-1-ZONE1	<i>standard</i>
		RAD2	C-TH-ZONE2	C-1-ZONE2	
		RAD3		C-1-ZONE3	
HORTOR	VERBE	RAD1	A-TH-ZONE1	D-1-ZONE1	<i>standard</i>
		RAD2	A-TH-ZONE2	D-1-ZONE2	
		RAD3		D-1-ZONE3	
LICET	VERBE	RAD1	E-TH-ZONE1	E-1-ZONE1	<i>impersonnel</i>
		RAD2	E-TH-ZONE2	E-1-ZONE2	
		RAD3		E-1-ZONE3	
ACCEDITUR	VERBE	RAD1	C-TH-ZONE1	D-1-ZONE1	<i>impersonnel, PTCP</i>
		RAD2	C-TH-ZONE2	D-1-ZONE2	
		RAD3		D-1-ZONE3	
INQUAM	VERBE	RAD1	I-TH-ZONE1	C-1-ZONE1	PRS.IPFV.IND, 3SG.PST.IPFV.IND, 2/3SG.FUT.IPFV.IND, 1/2.SG.PRS.PFV.IND
		RAD2	I-TH-ZONE2	C-1-ZONE2	
			I-TH-ZONE3	C-1-ZONE3	

TABLE 1.29 : Informations lexicales associées aux verbes latins

les niveaux suivants <sup>51</sup> :

Niveau 1 classes de radicaux ;

Niveau 2 thème ;

Niveau 3 exposants de temps, d'aspect et de mode ;

Niveau 4 marques de personne et de cas.

Les deux niveaux d'exposants comporteraient deux classes flexionnelles chacun. Ainsi les verbes appartenant dans l'analyse de ce chapitre aux classes A et C partageraient une classe flexionnelle au niveau 4, tout comme les verbes appartenant aux classes B et D. En revanche au niveau 3, ce seraient les verbes A et B qui partageraient une même classe flexionnelle, de même que les verbes appartenant aux classes C et D. La dérivation passive se ferait conjointement avec un changement de classe flexionnelle au seul niveau 4 mais maintiendrait les classes des niveaux 1 à 3 inclus. Cela éviterait notamment la redondance observable dans l'analyse donnée : en l'occurrence le processus dérivationnel de la passivation ne s'encombrerait plus de la correspondance systématique entre les classes A et B d'une part et C et D de l'autre. Une telle analyse mettrait également mieux en évidence la correspondance existante entre niveau thématique et classe flexionnelle de niveau 3 qui est indépendante du niveau 4.

## 1.5 Avantages et problèmes liés à la formalisation de la morphologie

Comme le montre la flexion automatique de notre lexique implémenté, notre analyse permet de rendre compte de l'intégralité des formes d'un lexique conséquent de verbes latins.

La prise en compte de l'intégralité des valeurs associées aux formes m-passives nous a menée à donner une analyse dérivationnelle du passif en latin et à poser des classes flexionnelles distinctes pour les verbes aux désinences m-actives et m-passives. Elle nous a également permis de rendre compte du fait que les verbes déponents ne sont pas les seuls verbes latins à afficher un paradigme avec des formes non finies surnuméraires puisqu'ils se comportent en cela comme les verbes que Flobert qualifie de *passiva tantum*.

Grâce à la factorisation de la représentation de la morphologie par des niveaux d'analyse indépendants et combinables (radicaux, thème et exposants), nous pouvons par ailleurs

---

51. Une telle analyse suppose dans le modèle que nous présentons au chapitre 4 un traitement différent pour les cases affectées par des règles porte-manteau dans l'analyse à trois niveaux de ce chapitre.



traiter de façon homogène les verbes à thème TH-A et TH-E d'une part et TH-I et TH-C de l'autre.

Ainsi, notre analyse ne repose que sur quatre classes flexionnelles, résumées dans le tableau 1.30.

CLASSE FLEXIONNELLE	THÈMES COMPATIBLES	DÉSINENCES
CLASSE A	TH-A/TH-E	<i>m-actif</i>
CLASSE B	TH-A/TH-E	<i>m-passif</i>
CLASSE C	TH-I/TH-C	<i>m-actif</i>
CLASSE D	TH-I/TH-C	<i>m-passif</i>

TABLE 1.30 : Les quatre classes flexionnelles du latin

La représentation par niveau d'analyse permet également d'identifier différents types d'hétéroclise qui peuvent apparaître dans les paradigmes. Nous avons ainsi été amenée à donner une définition explicite des partitions de paradigmes que constituent les zones (fussent-elles des zones de radicaux, des zones flexionnelles ou des zones thématiques). La définition formelle de ces notions est donnée au chapitre 4.

Enfin, notre implémentation complète nous a forcée à définir explicitement les types d'irrégularités, ponctuelles ou généralisées, qui apparaissent dans les systèmes morphologiques et à concevoir un système permettant de dissocier les irrégularités structurales (irrégularités dans la grammaire) des irrégularités ponctuelles (irrégularités associées à des entrées lexicales particulières). Il apparaît ainsi qu'une théorie de la morphologie apte à rendre compte de systèmes morphologiques complets ne peut se passer d'une théorie de l'interface entre lexique et grammaire, définissant notamment les informations constitutives d'une entrée lexicale.

Dans le chapitre suivant, nous présentons les concepts théoriques principaux intervenant dans la conception du modèle formel que nous introduirons dans la seconde partie de cette thèse. Nous y développerons en particulier la question de l'interface de la morphologie flexionnelle avec le lexique et présenterons en conclusion la vision modulaire articulée de la grammaire sur laquelle sera bâti notre modèle formel.

## Approches formelles de la morphologie flexionnelle

L'objectif de cette thèse est de proposer un modèle formel de la morphologie flexionnelle. Toutefois, le traitement de la morphologie comme composante indépendante de la grammaire, et notamment sa distinction par rapport à la syntaxe est encore aujourd'hui un sujet controversé. Outre la question de l'existence d'une morphologie autonome de la syntaxe et de la phonologie, se pose la question de la pertinence de l'opposition entre *morphologie constructionnelle* et *morphologie flexionnelle*. Dans ce chapitre nous présentons les débats théoriques autour de ces notions ainsi que les principaux arguments pour un traitement de la morphologie comme composante de la grammaire, distincte des autres et méritant un traitement autonome. Nous défendons en particulier le fait de traiter de façon séparée la morphologie flexionnelle en ce qu'un tel traitement autonome permet de rendre compte des caractéristiques idiosyncrasiques formelles apparaissant dans les paradigmes flexionnels.

Les approches existantes confondent également souvent morphologie et lexique. Dans ce chapitre, nous nous intéressons donc également au statut du lexique dans une approche qui accorde à la morphologie, et en particulier à la morphologie flexionnelle, l'entière importance qu'elle mérite au sein de l'architecture de la grammaire. Nous nous intéresserons enfin au statut de la composante de la morphologie flexionnelle par rapport au reste de la grammaire et proposerons la conception modulaire articulée de la grammaire sur laquelle est bâti le modèle formel de la morphologie flexionnelle que nous introduirons dans la seconde partie de cette thèse.

Nous développons ainsi à la section 2.1 les arguments qui justifient une approche

lexicaliste forte, lexématique, formellement explicite et réalisationnelle de la morphologie flexionnelle. Nous montrerons aussi l'avantage à accompagner cette approche d'une véritable théorie du lexique. À la section 2.3, nous décrivons également plus en détail deux modèles formels existants avec lesquels notre modèle partage le plus de points communs : *Paradigm Function Morphology* (Stump, 2001, 2006) et *Network Morphology* (Corbett & Fraser, 1993 ; Brown & Hippisley, 2012).

## 2.1 Délimitation du cadre conceptuel

### 2.1.1 Définition du lexicalisme

#### 2.1.1.1 Origine du lexicalisme

Dans la tradition générative, c'est depuis l'article de Chomsky (1970)<sup>1</sup> que l'idée que la morphologie constructionnelle comporte des caractéristiques différentes de celles de la syntaxe est largement admise. Chomsky défend en effet que la syntaxe ne peut rendre compte qu'exclusivement de correspondances régulières entre deux formes linguistiques et que les idiosyncrasies doivent être prises en charge par une composante externe à la grammaire appelée lexique. Compte tenu des idiosyncrasies observables en morphologie dérivationnelle, et notamment dans la dérivation permettant la création de noms déverbaux et déadjectivaux, Chomsky affirme que la morphologie dérivationnelle comporte une composante proprement lexicale qui la distingue des transformations nécessairement régulières de la grammaire, c'est-à-dire en fin de compte pour l'auteur de la syntaxe. Chomsky distingue ainsi la nominalisation gérondive en *-ing* en anglais, qui relève de la flexion, des autres types de nominalisation relevant d'une opération dérivationnelle. La différence cruciale entre ces deux types de nominalisation concerne leur prédictibilité sémantique, et en particulier celle de leur structure argumentale. Les gérondifs comme *refusing* « *refuser* » et *criticizing* « *critiquer* » prennent un objet direct et ont une structure argumentale homologue à celle du verbe correspondant (comparer les exemples (2.1-a) à (2.2-a) et (2.1-b) à (2.2-b)).

- 2.1 a. *John has refused the offer.*  
John a refusé la offre  
« *John a refusé l'offre.* »

---

1. *Remarks on Nominalization.*

- b. *John criticized the book.*  
 John a critiqué le livre  
 « *John a critiqué le livre.* »

- 2.2 a. *John's refusing the offer.*  
 John.GEN refuser.GRD la offre  
 « *Le refus de John de l'offre.* »

- b. *John's criticizing the book.*  
 John.GEN critiquer.GRD le livre  
 « *La critique du livre de la part de John* »

Les noms déverbaux ne prennent pas de complément direct, et leur structure argumentale est imprédictible : par exemple, *amusement* « *amusement* » a une structure argumentale rappelant celle d'un verbe passif, alors que *criticism* « *critique* » est homologue à un verbe actif, cf. les exemples 2.3 et 2.4 ci-dessous.

- 2.3 a. *John's amusement at the children's antics.*  
 John.GEN amusement à propos de les enfants.GEN pitreries  
 « *L'amusement de John quant aux pitreries des enfants* »

- b. *John's being amused at the children's antics.*  
 John.GEN être.GRD amusé à propos de les enfants.GEN pitreries  
 « *L'amusement de John quant aux pitreries des enfants* »

- c. #*John amuses the children's antics.*  
 John amuse les enfants.GEN pitreries  
 # « *John amuse les pitreries des enfants* »

- d. *The children's antics amuse John.*  
 les enfants.GEN pitreries amusent John  
 « *Les pitreries des enfants amusent John* »

2.4

- a. *John criticizes the book.*  
John critique le livre  
« *John fait la critique du livre.* »
  
- b. *John's criticism of the book.*  
John.GEN critique de le livre  
« *La critique du livre de la part de John* »

Le fait que ce changement de valeur soit hautement idiosyncrasique et dépende ainsi des unités lexicales soumises aux nominalisations en question a été proposé par Chomsky comme un argument indiquant que la morphologie dérivationnelle est une opération lexicale, externe à la syntaxe.

En revanche, ce type d'argumentation ne vaut que pour la morphologie constructionnelle. Comme nous l'avons noté dès le chapitre 1, la flexion est sémantiquement régulière, voir notamment (Wurzel, 1984 ; Boyé, 2011).

**Approches contemporaines non lexicalistes** Certaines approches contemporaines rejettent cependant l'argumentation de Chomsky (1970), et donc au lexicalisme.

Lieber (1992) propose un modèle qui comporte une définition assez complexe de la notion de lexique, élaborée dès (Lieber, 1981). Lieber prend en quelque sorte le contre-pied des conclusions de Chomsky (1970). Dans le modèle de Lieber, le lexique comporte l'ensemble des morphèmes (« séquences phonologiques représentant des unités minimales de sens ») de la langue. Ces morphèmes, habituellement considérés comme les éléments constitutifs des mots, sont dans le modèle de Lieber des éléments insérés depuis le lexique dans la structure sous-jacente de la syntaxe. Ils y sont combinés comme n'importe quel autre élément de la syntaxe et soumis aux mêmes opérations transformationnelles que celles employées par ailleurs dans les approches transformationnelles de type *Gouvernement et Liage*, *Government and Binding (GB)* (Chomsky, 1981). À la fin du processus de transformation, la grammaire est censée avoir généré, à partir de la structure sous-jacente, la structure de surface — des mots comme des phrases indifféremment.

La *Morphologie Distribuée (Distributed Morphology (DM))* introduite par Halle & Marantz (1993) se distingue de l'approche de Lieber (1992) en ce qu'elle ne suppose pas un lexique composé d'unités lexicales pleinement spécifiées, mais plutôt un ensemble d'unités sous-spécifiées en termes de phonologie mais représentant des combinaisons de

traits lexicaux et grammaticaux. Ces unités sous-spécifiées sont insérées en syntaxe et subissent, comme dans le cas de Lieber (1992), des opérations syntaxiques. Une fois les opérations syntaxiques effectuées, la composante phonologique insère les informations phonologiques jusqu'alors sous-spécifiées permettant la production concrète (*Spell Out* « épel ») de la phrase. Les idiosyncrasies sont gérées par une composante séparée, l'« encyclopédie » qui filtre les structures non conformes aux exigences lexicales.

### 2.1.1.2 Le lexicalisme fort

En partant du point de vue de Chomsky (1970) Bresnan (1982) soutient qu'il convient de considérer également les opérations flexionnelles comme des règles lexicales, dans la mesure où des formes fléchies peuvent parfois servir d'entrée à une autre règle lexicale. Elle donne l'exemple des adjectifs déverbaux anglais qui sont le résultat d'une conversion<sup>2</sup> à partir des participes passifs. Dans l'exemple 2.5, *unzipped* est ainsi le participe passif de UNZIP « ouvrir la fermeture éclair » ; mais *untouched* « intouché » ne peut être le participe passif, le verbe \*UNTOUCH ne faisant pas partie de la langue. Il ne peut donc s'agir que du dérivé de l'adjectif TOUCHED combiné au préfixe négatif *un-*.

2.5 (Bresnan, 1982, p. 29)

a. *The jacket was unzipped by someone wearing fingernail polish.*

Le blouson a été dézippé par quelqu'un portant du vernis à ongles

« *La fermeture éclair du blouson a été défaire par quelqu'un portant du vernis à ongles.* »

b. *The jacket was untouched by human hands.*

Le blouson a été intouché par humaines mains

« *Le blouson n'avait été touché d'aucune main d'homme.* »

2.6 (Bresnan, 1982, p. 29)

a. *Margaret's statement was considered profound.*

Margaret.GEN prise de position a été conçu comme profond

« *La prise de position de Margaret a été appréciée pour sa profondeur.* »

---

2. La conversion est une règle de construction au même titre que la dérivation, parfois également appelée *dérivation zéro*, cf. par exemple Kiparsky (1982). On considère généralement qu'elle se caractérise par l'égalité de forme entre le convers et une forme fléchie de la base et un changement de catégorie morphosyntaxique (Tribout, 2010).

- b. *That was a very **considered** statement.*  
Ce était une très réfléchie prise de position  
« *C'était une prise de position très réfléchie.* »

De même, dans l'exemple 2.6, le second *considered* ne peut être qu'un adjectif, le participe ne pouvant pas être modifié par l'adverbe *very*.

2.7 (Bresnan, 1982, p. 29)

- a. *The prisoners were **spared** execution.*  
Les prisonniers ont été épargné de exécution  
« *L'exécution a été épargnée aux prisonniers.* »
- b. *The **spared** prisoners were freed.*  
Les épargnés prisonniers ont été libérés  
« *Les prisonniers qui avaient été épargnés ont été libérés.* »

Enfin, dans l'exemple 2.7, le second *spared* ne peut être qu'un adjectif puisqu'il modifie le nom *prisoners* « *prisonniers* ».

Le fait que les adjectifs en question comportent la valeur du passif que n'a pas le lexème de base montre que la conversion n'a pu se faire à partir d'une forme active, mais comporte bien la valeur qu'a acquise la forme par la flexion passive. Ainsi dans l'exemple 2.7, *spared* « *épargné* » comporte bien la valeur passive que ne porte pas par défaut *SPARE* « *épargner* », cf. la phrase active correspondante dans l'exemple (2.8-a) *vs.* la phrase passive dans (2.8-b).

- 2.8 a. *The president **spared** the prisoners from execution.*  
Le président a épargné les prisonniers de l'exécution  
« *Le président a épargné l'exécution aux prisonniers.* »
- b. *The prisoners were **spared** execution (by the president).*  
Les prisonniers ont été épargnés exécution (par le président)  
« *Les prisonniers ont été épargnés de leur exécution (par le président).* »

Ainsi, si la forme fléchie peut servir d'entrée à des règles de création lexicale, la flexion ne peut être prise en charge par la syntaxe. La conséquence en est que la syntaxe ne manipule que des entrées lexicales complètes et entièrement fléchies. Les défenseurs du lexicalisme fort considèrent par conséquent que la dérivation et la flexion relèvent toutes deux de mécanismes également indépendants de la syntaxe.

Avec le recul du temps, l'argumentation de Bresnan ne semble pas totalement convaincante : l'exemple discuté dépend crucialement du passif dont on a vu au chapitre 1 que le statut flexionnel ou constructionnel n'allait pas de soi. Il reste que la stratégie du lexicalisme fort s'est avérée payante et a mis au jour des différences substantielles entre morphologie et syntaxe.

### 2.1.2 La morphologie autonome

Une conséquence du lexicalisme fort est la notion d'*autonomie de la morphologie* par rapport à la syntaxe. Cette conception est notamment défendue par Zwicky & Pullum (1988) qui en délimitant soigneusement les domaines de la phonologie et de la syntaxe définissent également le domaine de la morphologie, à la fois indépendante de l'une et de l'autre. Zwicky & Pullum, p. 263-264 montrent notamment qu'il y a des règles morphologiques directement dépendantes de la morphologie (ou de la morphosyntaxe) qui s'appliquent uniquement à des frontières morphologiques, contrairement aux règles phonologiques au sens strict qui s'appliquent au système phonologique dans son ensemble.

Aronoff (1994) montre par ailleurs qu'il y a des phénomènes qui ne peuvent être expliqués ni par la syntaxe, ni par la phonologie, mais qui sont de la morphologie pure : *Morphology by It-self*. Ces phénomènes recouvrent des cas où la morphologie affiche de la structure aucunement motivée par les autres composantes de la grammaire, voire contraire à ce que les autres composantes pourraient suggérer. Aronoff donne l'exemple des formes des participes du passé et du passif des verbes en anglais qui présentent systématiquement la même forme quand bien même rien dans leur fonction syntaxique ne motive une identité entre elles plutôt qu'avec d'autres formes. La fonction qui de façon régulière associe une forme commune à des traits morphosyntaxiques différents sans que rien en syntaxe ne motive cette égalité de forme est appelée *fonction morphomique* ou *morphome* par Aronoff (1994, p. 25).

Aronoff cite également l'exemple des radicaux latins. Les descriptions traditionnelles du latin (cf. chapitre 1) font état de trois radicaux verbaux distincts conditionnés par des traits morphosyntaxiques liés à des sous-paradigmes de leur système flexionnel. Ils



sont généralement qualifiés de radicaux du présent, du passé et du supin : RAD1 dans *amo* « j'aime », RAD2 dans *amāvī* « j'ai aimé » et RAD3 *amātum* « aimé ». Ces radicaux n'expriment pourtant pas dans leur répartition une réalisation morphologique de traits morphosyntaxiques régulière. Ainsi, le RAD3 peut également servir pour les formes du participe passé passif, le participe futur actif et pour les formes finies perfectives de la classe flexionnelle du passif<sup>3</sup>. Il ne semble pas y avoir de trait morphosyntaxique unique qui implique nécessairement l'emploi du RAD3. Cette distribution arbitraire, mais régulière d'un paradigme à l'autre, qu'affichent les radicaux du latin est décrite par Aronoff comme une structuration purement morphologique, indépendante des contraintes fonctionnelles de la syntaxe.

Corbett donne un exemple de structure morphologique pour les verbes du dhaasanac.

	SED « <i>marcher</i> »		YUUFUMI « <i>tousser</i> »	
	SG	PL	SG	PL
1.INCL	—	<i>seḏ</i>	—	<i>yuufumi</i>
1.EXCL	<i>seḏ</i>	<i>sieti</i>	<i>yuufumi</i>	<i>yuufeeni</i>
2	<i>sieti</i>	<i>sieti</i>	<i>yuufeeni</i>	<i>yuufeeni</i>
3.F	<i>sieti</i>	<i>seḏ</i>	<i>yuufeeni</i>	<i>yuufumi</i>
3.M	<i>seḏ</i>	<i>seḏ</i>	<i>yuufumi</i>	<i>yuufumi</i>

TABLE 2.1 : Morphème en dhaasanac (Corbett, 2010) (données de Tosco (2001))

Les verbes SED « *marcher* » et YUUFUMI « *tousser* » du tableau 2.1 présentent la même structure formelle, à savoir des syncrétismes entre 1SG.EXCL, 3SG.M, 1PL.INCL, 3PL.F et 3PL.M d'une part, et 2SG, 3SG.F, 1PL.EXCL et 2PL de l'autre, sans qu'aucune caractéristique phonologique ou syntaxique valable en dhaasanac ne requière ces regroupements.

L'existence d'objets purement morphologiques ne fait sens que si on admet l'autonomie de la morphologie ; l'argumentation d'Aronoff (1994) motive donc, par un tout autre biais que celui de Bresnan (1982), la nécessité d'une frontière nette entre morphologie flexionnelle et syntaxe.

## 2.1.3 Objet d'étude de la morphologie

### 2.1.3.1 Le débat entre *mots* et *morphèmes*

La définition de la morphologie sans doute perçue aujourd'hui comme la plus standard serait de dire que la morphologie étudie la construction des mots, cette construction

3. Cf. notre analyse au chapitre 1.

étant le mode d'agencement à l'intérieur du mot d'éléments signifiants minimaux appelés *morphèmes* (Baudouin de Courtenay, 1895).

L'utilité analytique de la notion de *morphème* est cependant discutée aujourd'hui au sein des différentes approches de la morphologie. Déjà Hockett (1954, p. 42) note que le morphème n'est pas un élément du langage mais constitue un « artefact de l'analyse » ou une « commodité pour la description ». Robins (1959, p. 128) défend une approche de la morphologie qui s'appuierait sur la notion de *mot* plutôt que de morphème. Son argument est que dans une analyse grammaticale, qui en soi constitue une abstraction à partir des phénomènes observés, le mot donne des indications plus fiables sur les relations grammaticales : autrement dit, il est plus immédiat d'abstraire à partir de mots complets qu'à partir de morphèmes individuels — puisqu'ils constituent eux-mêmes une abstraction différente. Matthews (1972, p. 124) indique par ailleurs que l'appariement entre forme et sens minimal constitutif de la notion de morphème ne résiste pas aux réalités observables.

Nous adopterons donc ici comme Matthews (1972, 1974), Anderson (1992) Aronoff (1994) et Stump (2001), une conception de la morphologie qui pose le mot, mais pas le morphème, comme primitive<sup>4</sup>. Nous emploierons une définition de mot qui s'appuie sur le modèle proposé par Matthews (1974, p. 64) et qui n'utilise donc pas la notion de morphème. Pour Matthews, un mot comme le verbe latin *amāmus* « nous aimons » est constitué d'une *racine* *am*, de *formatifs* optionnels comme la voyelle thématique *-a-* et éventuellement d'une marque flexionnelle appelée *exposant*, ici la désinence *-mus*. Un exposant est la réalisation phonétique d'une relation d'exponence, c'est-à-dire du marquage d'un certain ensemble de traits morphosyntaxiques.

### 2.1.3.2 La structure interne des mots

Une question distincte de la construction des mots est celle de la structure interne des mots. Pour les défenseurs d'un traitement syntaxique de la construction des mots, la structure interne des mots ressemble en général à des structures en constituants employées pour représenter la structure des phrases. Une telle approche est notamment défendue par Williams (1981). La structure des mots est alors en généralement considérée comme complètement transparente pour la syntaxe, de sorte que la syntaxe peut la modifier. Les

---

4. Les approches lexicalistes fortes, telles que nous les avons définies à la section 2.1.1, ne sont toutefois pas incompatibles avec les approches morphématiques de la morphologie. Ainsi la *Morphologie Constructionnelle* (*Construction Morphology*) de Booij (2010), certaines approches de la morphologie en LFG (*Grammaire Lexicale Fonctionnelle ou Lexical Functional Grammar* (Bresnan, 2001 ; Dalrymple, 2001)), notamment (Nordlinger, 1998 ; Nordlinger & Sadler, 2004), ou les travaux de Paul Kiparsky, notamment (Kiparsky, 1982), constituent des approches lexicalistes employant la notion de morphème.

modèles de Lieber (1992) et Marantz (1997) relèvent de ce type d'approches.

À l'opposé de cela se situent les approches qui défendent l'idée selon laquelle les mots sont construits à partir de procédés complètement différents de ceux employés pour la construction des phrases en syntaxe et qu'aucune structure interne des mots ne peut être accessible à la syntaxe (Anderson, 1992 ; Stump, 2001). En particulier, ces approches considèrent généralement qu'il n'y a pas dans la construction des mots de relations tête-dépendant. Quelques exceptions ponctuelles à cette règle générale sont cependant citées par Stump (2001, chapitre 4). L'auteur montre qu'un certain nombre d'opérations constructionnelles endocentriques préserve certaines propriétés morphosyntaxiques de la base. Il cite ainsi les diminutifs péjoratifs en *-iška* du russe et les diminutifs bretons en *-ig* qui préservent le genre de leur base<sup>5</sup>, ainsi que les diminutifs en *-aka* du barasano du sud<sup>6</sup> pouvant se combiner à des mots de catégories syntaxiques diverses mais préservant la catégorie de la base.

Ces approches rejoignent par ailleurs également le PRINCIPE D'INTÉGRITÉ LEXICALE initialement formulé par Bresnan et décrit plus en détail dans (Bresnan & Mchombo, 1995, p. 187). Il stipule que les éléments manipulés par la syntaxe sont des mots morphologiquement complets, c'est-à-dire des mots-formes entièrement construits. Toutefois, Anderson (1992, p. 24) montre sur l'exemple de l'esquimo qu'il peut y avoir des exceptions au principe d'intégrité lexicale. L'auteur y montre qu'en esquimo la structure syntagmatique en syntaxe ne recoupe par nécessairement la structure en mots observable en morphologie. Des phénomènes d'incorporation nominale dans le verbe y ont pour effet que la tête de certains syntagmes nominaux objets se trouve morphologiquement liée au verbe. La structure morphologique n'est ainsi pas compatible avec la structure syntaxique.

#### 2.1.4 Une vision réalisationnelle de la morphologie

Les approches de la morphologie se répartissent principalement en deux conceptions de la formation des mots : les APPROCHES INCRÉMENTALES et les approches APPROCHES RÉALISATIONNELLES.

Parmi les premières, il y a notamment le modèle proposé par Steele (1995). La MORPHOLOGIE ARTICULÉE (*Articulated Morphology (AM)*) (Steele, 1995) est une approche incrémentale non morphématique : chaque opération, de construction ou de flexion, a pour effet de spécifier un peu plus le sens, respectivement lexical et grammatical,

---

5. Contrairement, par exemple, aux diminutifs allemands en *-chen* qui assignent systématiquement le genre neutre au nom dérivé.

6. Langue tucanoane parlée en Amazonie, en Colombie.

d'un mot donné. En particulier, plusieurs opérations de flexion peuvent être appliquées les unes après les autres, spécifiant à chaque fois un peu plus les traits morphosyntaxiques portés par le résultat de l'opération par rapport à l'entrée. Les règles de flexion sont ainsi organisées en blocs de règles appliquées séquentiellement. Steele propose 1) que ces blocs soient de véritables étapes dans la construction du sens de sorte que le sens du mot reste sous-spécifié jusqu'au résultat de la toute dernière règle et que 2) les opérations soient sensibles au sens apporté par les opérations précédentes, créant ainsi une structure motivée sur laquelle s'appuient les règles subséquentes. L'argument de Steele est que l'ordre dans lequel s'appliquent les différentes règles de réalisation n'est pas arbitraire, mais constitue une incrémentation cohérente du sens d'un mot donné. Une règle d'un bloc ultérieur ne pourra s'appliquer que si le mot en question a déjà été spécifié dans les blocs antérieurs pour le domaine d'application de la règle en question. Nous illustrons cette idée à partir de deux exemples. L'exemple 2.9 du français illustre le fait que l'application d'une règle réalisant la personne sur un mot avec une base adjectivale impose d'abord l'introduction d'un suffixe dérivatif par une autre règle construisant un radical verbal déadjectival sur cette base. De même, dans l'exemple 2.10, l'application des marques personnelles du parfait en latin impose l'introduction préalable de la marque du parfait. Elles ne peuvent s'appliquer après la règle introduisant l'imparfait illustrée en (2.10-b).

2.9      sêpl-if-(j)ð  
           simple-VERBE-1PL

2.10    a.    *amav-i-sti*  
           aimer.RAD2-PRS.PFV.IND-2SG.PRS.PFV.IND  
           « *tu as aimé* »

          b.    *am-a-ba- \*sti*  
           aimer-RAD1-VOYELLE THÉMATIQUE-PST.IPFV.IND- \*2SG.PRS.PFV.IND

Il existe toutefois des arguments importants allant à l'encontre d'une structuration interne des mots. Le plus connu concerne l'exponence multiple : un mot peut aisément comporter plusieurs manifestations d'un même trait morphosyntaxique. Prenons l'exemple du participe passé en allemand. Le participe passé en allemand est généralement marqué par un préfixe *ge* (*ge*) et un suffixe *-t* (*t*), cf. l'exemple 2.11. Pour les verbes fléchis selon cette règle régulière, les deux suffixes portent tous les deux la valeur de participe passé (PST.PTCP),

l'un étant insuffisant sans l'autre.

- 2.11    *Heike*            *hat*                            *Lukas*            *ein*            *Eis*  
           Heike.[NOM]    AUX.3SG.PRS.IND    Lukas[DAT]    un.ACC.SG.N    glace.ACC.SG.N  
           *ge-kauf-t.*  
           PST.PTCP-acheter-PST.PTCP  
           « *Heike a acheté une glace à Lukas.* »

Une autre exemple est celui de la forme *iront* en français (exemple (2.12-a)). Le trait FUTUR est exprimé dès le choix du radical supplétif *i-*. Il s'exprime ensuite de façon redondante par le suffixe temporel *-r-*, puis encore par l'exposant fusionnel *-ont*. Dans la mesure où le trait FUTUR est ainsi déjà exprimé dès le choix du radical, on ne peut pas dire que *-i-* ni *-ont* incrémentent la valeur du FUTUR dans *iront*; en revanche, ils le font dans *fini-r-ont* à l'exemple (2.12-b).

- 2.12    a.    *i-r-ont*  
           aller.FUT-FUT-3PL.FUT  
        b.    *fini-r-ont*  
           finir.-FUT-3PL.FUT

À première vue, ce système ressemble aux blocs de règles utilisés dans le formalisme d'Anderson (1992) et en PFM Stump (2001) qui pourtant récusent dans les cas général<sup>7</sup> l'idée de structure interne des mots. La Morphologie Articulée se distingue des approches d'Anderson ou Stump en ce que ces derniers 1) considèrent que les règles de réalisation réalisent des traits déjà pleinement spécifiés pour le mot qu'elles forment et 2) que l'ordonnancement par blocs de ces règles ne correspond qu'à une hiérarchisation externe des réalisations phonologiques des traits en question et non pas à une nécessité interne liée à l'incrémentation.

Dans la mesure où les approches réalisationnelles partent du principe que les traits réalisés par un mot-forme sont déjà pleinement spécifiés avant même l'application des règles de réalisation de la morphologie (qui se contentent de les exprimer par des exposants), elles ne contraignent en rien les cas d'exponence multiple. Ce fait nous semble approprié dans la mesure où l'exponence multiple est fréquemment observable à travers les langues

---

7. C'est-à-dire en dehors des cas particuliers cités ci-dessus.

et peut même parfois prendre des dimensions considérables : Harris (2009) montre ainsi qu'en batsbi<sup>8</sup> le marquage de genre et de nombre peut se manifester jusqu'à cinq fois sur le même verbe. Une telle phénoménologie, pourtant observable, serait prédite comme étant impossible par des approches incrémentales. Cet exemple nous conforte ainsi dans le point de vue que ce n'est pas le rôle de la théorie de formuler des contraintes *a priori* sur les phénomènes morphologiques observables mais que ces contraintes éventuelles doivent découler de l'étude typologique des systèmes morphologiques existants.

## 2.1.5 Flexion et construction et ce qui les distingue

Nous avons vu à la section 2.1.1.2 les arguments plaidant pour un traitement de la flexion et de la morphologie constructionnelle indépendant de celui de la syntaxe.

Le regroupement de construction et flexion dans un dispositif grammatical commun ne signifie pour autant pas que les deux types d'opérations morphologiques sont de même nature.

Dans une approche morphématique incrémentale comme celle de Lieber (1992) ou Booij (2010), il semble naturel de considérer que toutes les opérations morphologiques ont en commun d'augmenter la quantité d'information véhiculée par un mot donné. Ainsi la combinaison du morphème *pwax* avec le morphème *je* augmente l'information « dénotation d'un fruit X », où X est une poire, de l'information « arbre sur lequel pousse le fruit X ». De même, nous pouvons considérer que, en simplifiant (beaucoup), fléchir le verbe MANGER pour la forme 1PL.PRS.IPFV.IND.ACT *mangeons* augmente l'information contenue dans le sens lexical de MANGER de l'information que le mangeur est une entité multiple dont fait partie le locuteur et que l'action de manger est un processus en cours au moment de l'énoncé.

*Split Morphology* Anderson (1982, p.587), en revanche, définit la morphologie flexionnelle comme la partie de la morphologie qui joue un rôle en syntaxe. Il s'appuie pour cela sur le concept de « morphologie scindée » (*Split Morphology*) également défendu par Perlmutter (1988) dans un cadre lexicaliste faible. Les arguments pour la scission entre morphologie dérivationnelle et morphologie flexionnelle reposent essentiellement sur l'idée que la morphologie flexionnelle constitue une interface avec la syntaxe et répond à des contraintes imposées par cette dernière. Or, dans une approche générative classique et lexicaliste faible, la syntaxe agit sur des éléments du lexique, qui sont éventuellement les résultats d'opérations morphologiques constructionnelles. Les approches lexicalistes faibles dans la tradi-

---

8. Langue nakh-daghestanienne parlée en Géorgie.

tion de Chomsky (1970) admettent une séparation nette entre le lexique, dépositaire des idiosyncrasies de la langues, et la syntaxe, domaine des transformations régulières. La morphologie flexionnelle étant conditionnée par des contraintes syntaxiques, elle ne peut être conçue autrement que distincte de la morphologie constructionnelle. Anderson indique par ailleurs comme argument en ce sens que, dans la construction d'un mot, l'affixation relevant de la morphologie flexionnelle est toujours réalisée à l'extérieur de l'affixation dérivationnelle<sup>9</sup>. Ce fait est également mis en rapport avec l'affirmation que le résultat de la morphologie flexionnelle ne peut servir d'entrée à des opérations constructionnelles.

**Remise en cause de la *Split Morphology*** Ce dernier point est contesté par Booij (1993, 1996). Nous avons d'ailleurs montré à la section 2.1.1.2 que c'est justement le fait que le résultat de la flexion puisse servir d'entrée à des opérations de conversion pour les adjectifs déverbaux qui a mené Bresnan (1982) à poser les principes d'un lexicalisme fort, rassemblant dans un même domaine extra-syntaxique la morphologie constructionnelle et flexionnelle.

De même, Booij s'appuie sur le fonctionnement des diminutifs et du pluriel en néerlandais (qui tous deux peuvent servir d'entrée à des règles de construction) pour défendre une vision qui traiterait ensemble, et en dehors de la syntaxe, la flexion et la construction. Il distingue cependant entre deux types de flexion, la *flexion inhérente* d'une part et la *flexion contextuelle* de l'autre. Ainsi la flexion inhérente concerne la partie de la flexion qui produit un effet syntaxique et/ou sémantique, comme le pluriel d'un nom qui dénote souvent plus d'un des individus dénotés par le nom en question. La flexion contextuelle comprend les phénomènes d'accord et de rection. Ainsi, dans l'exemple 2.13, le nom FILLE fléchi au pluriel dénote effectivement plus d'une fille. Le fait que *petit* soit également au pluriel et au féminin ne modifie aucunement la qualité d'être petit, mais indique uniquement que cette qualité s'applique à *filles*.

### 2.13 les petites filles

Selon les arguments de Booij (1993, 1996), la morphologie n'établirait pas de scission entre les opérations flexionnelles et dérivationnelles. En revanche, si la morphologie traite bien de la morphologie constructionnelle et de la flexion inhérente, Booij considère que la flexion contextuelle relève de la syntaxe.

---

9. On sait depuis que cette affirmation n'est en fait pas vérifiée, cf. par exemple Booij (1993).

**Différencier flexion et dérivation dans un cadre lexicaliste fort** D'autres auteurs ont depuis avancé des arguments pour différencier la flexion et la dérivation, et ce même dans un cadre lexicaliste fort, cf. (Matthews, 1972, 1974 ; Anderson, 1992).

Anderson (1992, p. 75-77) indique sept raisons pour lesquelles il lui semble pertinent de différencier la morphologie flexionnelle de la morphologie constructionnelle, et en particulier de la dérivation, et ce même dans une approche lexicaliste forte.

1. Les règles flexionnelles présentent souvent des réalisations de type PORTE-MANTEAU, c'est-à-dire qu'un morphe donné peut réaliser plus d'une valeur en même temps. Ainsi la marque *o* dans la forme *nɔʁmo* de l'adjectif *NORMAL* marque à la fois le masculin et le pluriel. Les règles constructionnelles ne présentent en général pas ce type de marquage fusionné.
2. On ne trouve pas de morphe qui serve à la fois à la construction et à la flexion.
3. La dérivation et la flexion sont ordonnées quand les deux sont présentes dans une langue. Si dérivation et flexion sont réalisées par de l'affixation, les affixes dérivationnels apparaîtront plus proches de la racine, les affixes flexionnels seront situés à la périphérie.
4. Certains patients atteints d'aphasie peuvent perdre la maîtrise de la morphologie flexionnelle et de la syntaxe quand bien même ils restent compétents en morphologie constructionnelle. Il cite notamment les travaux de Micelli & Caramazza (1989) et Badecker & Caramazza (1989)<sup>10</sup>.
5. Les règles flexionnelles, contrairement aux règles dérivationnelles (et en général aux règles de construction), sont pleinement productives, tandis que les règles de construction présentent souvent des trous dus à des phénomènes de *blocage*. Parmi ces blocages, certains sont dus au fait que la forme cible est déjà occupée par un homonyme, cf. la dérivation régulière *BOULANGER/BOULANGÈRE* *vs.* le blocage *CAFETIER/\*CAFETIÈRE* où *cafetière* désigne déjà « une machine à faire du café ». On trouve aussi d'autres blocages, plus idiosyncrasiques encore, cf. les noms déverbaux en *-aison* comme *LIVRER/LIVRAISON*, *PENDRE/PENDAISON*, *DÉMANGER/DÉMANGEAISON*, mais *MANGER/\*MANGEAISON*. Le nom *\*MANGEAISON* n'est pas « occupé » dans le lexique et le fait que *DÉMANGEAISON* existe montre par ailleurs que ce ne sont pas non plus des contraintes phonologiques qui conditionnent ce blocage. Il s'agit d'une simple idiosyncrasie, propre à la morphologie constructionnelle.

---

10. Cet argument est au fond un argument fort pour les approches lexicalistes faibles défendant un traitement homogène de la morphologie flexionnelle et de la syntaxe dans une même composante de la grammaire.



6. La construction change souvent la catégorie morphosyntaxique à laquelle appartient un mot, la flexion opère sans changer la classe, cf. *manger/mange/mangeons* qui sont tous des formes verbales *vs.* la construction de noms déverbaux sur le patron LIVRER/LIVREUR.
7. Le sens apporté par une règle flexionnelle est pleinement prédictible à partir du sens lexical de l'unité lexicale fléchie et des traits morphosyntaxiques exprimés par la flexion. Les règles de construction font souvent l'objet de glissements de sens importants, cf. aussi Wurzel (1984, p. 36) cité au chapitre 1.

Avec le recul du temps, chacune de ces observations devrait sans doute être nuancée et interprétée comme caractérisant des distinctions typiques entre flexion et dérivation, mais pas comme des oppositions entièrement catégoriques.

Ainsi, on trouve des suffixes dérivationnels de type porte-manteau dans les systèmes flexionnels. En français, par exemple, les suffixes dérivationnels dé-adjectivaux *-eur* et *-rice* dénotent à la fois un nom d'agent et le genre<sup>11</sup>.

Aronoff (1994, pp. 126-127) montre à l'exemple du système des binyan<sup>12</sup> du hébreu que la distinction entre flexion et dérivation n'est pas toujours évidente. Les procédés permettant de changer de binyan comportent des caractéristiques qui par certains côtés ressemblent à de la flexion<sup>13</sup> et par d'autres s'apparentent plus à de la flexion<sup>14</sup>.

L'argument selon lequel la flexion est réalisée après la construction a cependant été réfuté pour un certain nombre d'exemples, notamment pour le marquage du pluriel des noms néerlandais servant de base à la construction d'un autre nom singulier (Booij, 1993) ou des composés allemands à partir d'un nom au pluriel.

La morphologie supplétive produit des effets similaires au blocage. Nous en trouvons un exemple dans la forme du verbe ÊTRE au 1SG.PRS.IND *nous sommes* en français qui bloque la réalisation de *nous \*étons*.

Certaines formes fléchies comme les participes ou les gérondifs affichent des propriétés mixtes. Ainsi, le gérondif en anglais, tout en étant généralement considéré comme une forme verbale fléchie, a-t-il des propriétés verbales (il peut prendre un objet direct) et des propriétés nominales (il peut être modifié par un SN au génitif), cf. l'exemple 2.14 ci-dessous. De même, les participes, tout en étant des formes verbales, ont-ils souvent des propriétés adjectivales.

---

11. Masculin pour *-eur* et féminin pour *-rice*.

12. Les binyan constituent le classement traditionnel des verbes sémitiques en classe de conjugaison.

13. Glissements sémantiques, trous arbitraires.

14. Structuration des formes en paradigmes.

- 2.14 John's remembering his school  
Jean.GEN souvenir.GRV son école  
« le fait que Jean se souvienne de son école »

Enfin, la question de la prédictibilité du sens semble être une question de degré plutôt qu'un critère absolu, cf. les sens associés au passif en latin décrits à la section 1.1.2.2 du chapitre 1.

**La notion de lexème** Dans la tradition initiée par Matthews (1972, 1974), la différentiation entre la morphologie constructionnelle et la morphologie flexionnelle s'appuie sur la définition de la notion de *lexème*.

Un lexème selon Matthews (1974, p. 26) est une abstraction de l'ensemble des formes fléchies, ou *mots forme*, d'une entrée lexicale donnée : un lexème particulier comporte un sens lexical particulier ainsi que l'information permettant d'identifier l'intégralité de ses formes, c'est-à-dire de son *paradigme flexionnel*.

Pour Aronoff (1994, p. 11) le lexème est un membre (réel ou potentiel) d'une catégorie lexicale majeure. Il possède à la fois une forme et un sens mais il n'est assimilable à aucun des deux. Un lexème existe indépendamment de tout contexte syntaxique particulier. Un *mot grammatical* (ou *mot forme* chez Matthews) correspond à un lexème donné dans un contexte syntaxique particulier, c'est-à-dire lorsqu'il est assorti de traits morphosyntaxiques spécifiques, et morphologiquement réalisé en fonction de ces traits. Les mots grammaticaux sont les membres des paradigmes d'un lexème particulier<sup>15</sup>.

La notion de lexème fonde également ce que Fradin (2003) appelle les *approches lexématiques*. Les approches lexématiques différencient ainsi les opérations de construction lexicale (*Lexical Formation Rules (LFR)*) des opérations flexionnelles ou règles de réalisation des formes particulières du paradigme d'un lexème donné de l'autre.

## 2.1.6 La notion de paradigme

La notion de paradigme sous-jacente à la notion de lexème décrite à la section 2.1.5 a cependant une origine bien plus ancienne. Hockett (1954, p. 396) souligne qu'elle soutient ce qu'il appelle l'approche dite *Mot et Paradigme (WP)*, le modèle utilisé par les grammairiens anciens, notamment pour les descriptions traditionnelles du latin ou du grec,

---

15. Dans la suite de cette thèse, nous emploierons le terme de *mot forme*, ou même plus simplement de *mot*.

dans la grammaire sanskrite de Pāṇini (env. 500 AEC) et dans bon nombre de descriptions de langues plus modernes.

Le modèle WP, un moment abandonné au profit des approches morphématiques dont Hockett (1954) fait la synthèse magistrale, a connu un renouveau théorique par les travaux de Robins (1959), puis Matthews (1972, 1974); Anderson (1982); Stump (2001, 2006) et les approches de la MORPHOLOGIE IMPLICATIONNELLE (Blevins, 2006). Il repose sur le constat que la morphologie autonome (telle qu'elle est définie à la section 2.1.1.2) conditionne crucialement la structure des paradigmes et notamment les rapports existant entre les cases des paradigmes (cf. les syncrétismes observables dans les cases identiques des paradigmes du dhaasanac décrits à la section 2.1.5). Dans une approche lexématique, l'étude de la structure des paradigmes est ainsi un des enjeux principaux de l'étude de la morphologie flexionnelle.

## 2.1.7 Traitement de l'allomorphie de radical

**2.1.7.0.1 Allomorphie au sein des approches Mot et Paradigme** Dans les approches WP, les auteurs ont souvent cherché à identifier ce qui dans les cases d'un paradigme représente le matériau phonologique réalisant la flexion, c'est-à-dire l'exponence au sens de Matthews (1974), et ce qui correspond au matériau phonologique commun à toutes les cases (ou du moins un nombre significatif d'entre elles). Cette partie commune est appelée *radical* dans le modèle de Matthews (1974, p. 64). Prenons par exemple le paradigme du verbe MANGER au tableau 2.2<sup>16</sup>. Nous pouvons y distinguer le radical, constant à travers toutes les cases, mǎꝝ combiné à chaque fois à l'exposant (souvent nul) approprié pour la case concernée.

	SINGULIER	PLURIEL
P1	mǎꝝ	mǎꝝ-ḥ
P2	mǎꝝ	mǎꝝ-e
P3	mǎꝝ	mǎꝝ

TABLE 2.2 : Radical et exposants dans le paradigme du verbe MANGER au présent de l'indicatif

Parfois, comme dans le cas de BOIRE, l'identification de la partie commune est plus difficile. On dira que certaines cases présentent des radicaux différents du radical attendu.

16. Ce découpage correspond au découpage traditionnel donné pour les verbes du français et décrit notamment dans le *Bescherelle* (Arrivé, 1997).

	SINGULIER	PLURIEL
p1	bwa	byv- $\tilde{s}$
p2	bwa	byv-e
p3	bwa	bwa-v

TABLE 2.3 : Allomorphie du radical dans le paradigme du verbe *boire* au présent de l'indicatif

Dans le cas de BOIRE, l'alternance entre les radicaux *bwa* et *byv* constitue une propriété du paradigme de BOIRE, et ce plus encore que les caractéristiques des formes individuelles. Ainsi, l'allomorphie des radicaux donne aussi des indications sur le système morphologique dans son ensemble.

Pour Aronoff (1994, chap. 2), l'allomorphie de radicaux latins constitue un cas de structure morphomique. Un même radical peut apparaître dans une collection morphosyntaxiquement arbitraire de cases. L'inventaire des radicaux morphologiques est appelé *partitions* chez Pirrelli & Battista (2000) et *espaces thématiques* par Bonami & Boyé (2003, 2006). On retrouve la même idée chez Anderson (1992), Maiden (1992) et Brown & Hippisley (2012) (voir Bonami (2012) pour un panorama des approches à l'allomorphie).

Malgré les avantages d'un traitement explicite de l'allomorphie radicale, notamment pour la caractérisation de la structure des paradigmes, Blevins (2006) note que la segmentation d'un mot en radical et exposant n'est pas toujours immédiate. Faire un choix plutôt qu'un autre influence largement l'analyse dans son ensemble. Les données du tableau 2.3 sont par exemple découpées différemment dans Bonami & Boyé (2003), comme illustré par le tableau 2.4. Nous constatons que la réalisation de l'exponence n'est pas la même pour MANGER et BOIRE dans l'analyse traditionnelle indiquée dans les tableaux 2.2 et 2.3. Bonami & Boyé proposent une analyse des verbes français qui fait intervenir les mêmes exposants pour tous les verbes. Ainsi, le paradigme de BOIRE serait découpé comme dans le tableau 2.4 dans leur analyse.

	SINGULIER	PLURIEL
p1	bwa	byv- $\tilde{s}$
p2	bwa	byv-e
p3	bwa	bwav

TABLE 2.4 : Allomorphie du radical dans le paradigme du verbe *boire* selon Bonami & Boyé (2003)

Une autre analyse possible serait celle illustrée par le tableau 2.5 qui tenterait de

maintenir un radical identique pour toutes les cases du paradigme de BOIRE.

	SINGULIER	PLURIEL
p1	b-wa	b-yvõ
p2	b-wa	b-yve
p3	b-wa	b-wav

TABLE 2.5 : Le paradigme du verbe *boire* sans alternance de radicaux

Compte tenu de l’ambiguïté sous-jacente aux découpages entre radical et exposants, Blevins défend une approche de la morphologie qui abandonne la présupposition d’un découpage nécessaire des mots et n’accorde pas de place centrale à l’identification des radicaux.

Dans le modèle que nous proposons au chapitre 4, nous employons la notion de radical. Mais nous montrerons au chapitre 6 que l’arbitraire peut être contenu, voire éliminé, si l’on s’appuie sur une comparaison explicite des conséquences des choix de découpage en termes de compacité de l’analyse.

### 2.1.8 La structure du lexique

L’opposition entre grammaire et lexique intervient de manière récurrente dans les pages qui précèdent. Dans cette section nous présentons différentes manières dont l’articulation entre ces deux composantes a pu être envisagée.

Le lexique, cet « ailleurs » qui n’est pas la grammaire La plupart du temps, les théories développées ne se soucient que marginalement de la question du rapport entre lexique et grammaire. Le traitement du lexique proposé dans un des travaux fondateurs de la morphologie en grammaire générative, le modèle proposé par Halle (1973) en est une illustration particulièrement frappante. Le modèle de Halle s’appuie sur les constats faits par Chomsky (1970) à propos de la nominalisation en anglais. Alors même qu’il prend pour point de départ les remarques de Chomsky pour poser la question de ce qui constitue l’inventaire des mots existants dans la langue et de ce qui permet de définir et prendre en compte les caractéristiques idiosyncrasiques individuelles de certains mots, sa tentative de réponse ne se solde que par l’évacuation de ladite question.

Sa réponse se résume par le diagramme illustré par la figure 2.1. Dans un processus génératif représenté par une suite de boîtes correspondant aux différentes composantes de la grammaire, des listes de morphèmes sont en un premier temps traitées par des règles de

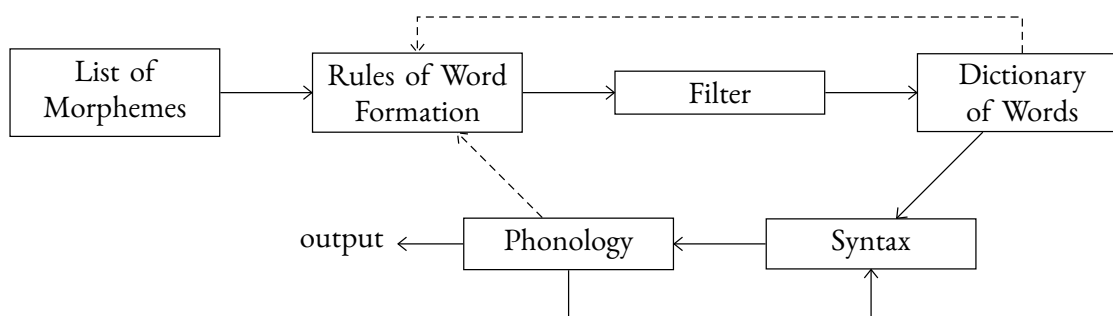


FIGURE 2.1 : Architecture de la grammaire selon le modèle de Halle (1973)

construction de mots qui génèrent toutes les formes possibles. Puis elles sont filtrées par un filtre indépendant qui supprime toutes les formes surgénérées ne figurant pas effectivement dans la langue. Les résultats filtrés sont ensuite rassemblés dans le dictionnaire qui fournit les éléments qui seront traités par la syntaxe, puis par la composante phonologique. Dans la mesure où certaines règles de construction de mots peuvent s'appliquer sur des mots déjà construits, les mots construits peuvent également servir d'entrées à ces règles (ce qui est représenté par la flèche en pointillée dans le schéma). Par ailleurs, Halle constate que certaines règles de construction sont contraintes par la forme résultant de règles phonologiques appliquées au résultat de transformations syntaxiques. Il propose donc également une possibilité pour les résultats de règles phonologiques de servir d'entrée à de nouvelles règles de construction de mots.

Le point qui nous intéresse particulièrement ici est le fait que quand bien même Halle (1973) pose la question du stockage des unités lexicales de la langue, il ne répond que par l'introduction d'une composante de filtre tout-puissant qui de façon arbitraire et obscure opère un tri sur les formes générables. En fin de compte, sa proposition ne revient qu'à masquer un problème pourtant dûment identifié.

**Les travaux du Lexique Grammaire** Ce traitement obscur réservé au lexique dans le modèle de Halle contraste avec les travaux entamés à la même époque dans le cadre théorique du *Lexique Grammaire* (Gross, 1975) et qui ont une influence considérable sur la communauté française. Ils sont essentiellement connus pour la constitution de tables dans lesquelles ont été compilées de riches informations (essentiellement syntaxiques) sur de larges quantités d'entrées lexicales. Ainsi, les tables du Lexique Grammaire comportent des listes de propriétés considérées comme lexicales pour chaque entrée, stipulant notamment la catégorie syntaxique, les cadres de sous-catégorisation, les transformations<sup>17</sup> possibles,

17. C'est-à-dire les changements de cadres de sous-catégorisation possibles.

*etc.* . Ce type d'informations varie en effet d'une entrée à l'autre. En français, par exemple, il existe deux homonymes VOLER<sub>1</sub> et VOLER<sub>2</sub> qui outre le fait que le premier désigne « le fait de se déplacer dans l'air » et le second « le fait de dérober un objet » se distinguent par le fait que le premier soit intransitif et le second transitif. Cette différence est purement lexicale et ne pourrait aucunement être déduite de la forme des lexèmes en question. Quand bien même les travaux du Lexique Grammaire concernent essentiellement la syntaxe, il existe également quelques tables spécifiques aux phénomènes de dérivation, telles les tables concernant les verbes causatifs déadjectivaux comme AFFAIBLIR (table 32RA, (Boons et al., 1976)).

Par ailleurs, les travaux du Lexique Grammaire comportent aussi une théorie de la morphologie flexionnelle, quoique moins connue et moins centrale dans l'ensemble de ces travaux. Le modèle de Courtois (1995) développé dans le cadre du Lexique Grammaire est un modèle lexématique reposant sur l'idée que chaque lexème (entrée lexicale dans une table) appartient à une classe flexionnelle représentée formellement par une catégorie morphosyntaxique et un numéro d'identification. Chaque catégorie morphosyntaxique est de son côté associée à un ensemble de structures morphosyntaxiques qu'elle peut réaliser. Chaque classe flexionnelle liste alors les exposants correspondant aux combinaisons de traits morphosyntaxiques définies pour la catégorie en question, permettant ainsi de construire pour chaque lexème l'ensemble des formes de son paradigme. Le fait d'associer chaque lexème à une classe flexionnelle d'une part, et de définir indépendamment les exposants réalisés par chaque classe flexionnelle de l'autre constitue une définition de la morphologie flexionnelle pleinement autonome du lexique dans sa composante grammaticale et une définition explicite de ce qui constitue une entrée lexicale pour la morphologie. De ce point de vue, le Lexique Grammaire constitue une véritable exception parmi les théories grammaticales contemporaines.

**Le lexique comme dépôt des irrégularités de la langue** En dehors de travaux du Lexique Grammaire, le lexique reste dans la plupart des théories un simple dépôt pour les irrégularités de la langue. Cette vision a son origine dans les propositions de Bloomfield (1933, 1939). L'approche bloomfieldienne consiste à faire du lexique une collection d'exceptions qui ne sont pas descriptibles par les règles de la grammaire supposée régulière. Cette vision est également partagée par des approches plus modernes, cf. Di Sciullo & Williams (1987).

L'idée d'un traitement des irrégularités de la langue par une composante appelée lexique est également à l'origine du développement du lexicalisme au sens de Chomsky

(1970) puisque c'est justement le constat que la dérivation s'accompagne d'idiosyncrasies importantes contraires à la régularité traitée par les règles de la syntaxe qui a mené l'auteur à proposer un modèle dans lequel (une partie de) la construction des mots serait évacuée dans la composante lexicale de la langue, fondamentalement distincte de la grammaire. Le rôle de la grammaire s'arrête ainsi à la mise en œuvre de transformations régulières appliquées aux éléments produits par le lexique, tandis que le lexique fait office de boîte dans laquelle on stocke les phénomènes non décrits.

**Morphologie autonome dans le lexique** Dans les approches lexicalistes fortes, les opérations morphologiques sont toutes traitées (de façon plus ou moins explicite) dans le lexique. Pour LFG, Bresnan (1982) propose des règles lexicales autant pour la construction que pour la flexion. Il s'agit de règles établissant des relations entre des éléments du lexique. Le traitement de ces relations est ainsi considéré comme fondamentalement interne à la composante lexicale de la langue.

Cependant, la plupart des travaux dans ces formalismes ne s'intéressent que marginalement aux questions de la morphologie. Dans Bresnan (1982), l'enjeu majeur semble plutôt être de délimiter ce qui dans une grammaire LFG sépare le lexique de la syntaxe, et non une réflexion sur la morphologie en tant que telle. Parmi les études les plus poussées sur le statut de la morphologie dans les grammaires d'unification figurent sans doute les travaux de Rachel Nordlinger (1998), pour LFG et le modèle de Jean-Pierre Koenig (1999) en HPSG.

Notons enfin ici que dans certains formalismes, la morphologie appartient à une composante morphologique indépendante, mais à l'intérieur même du lexique — au lieu d'être une partie autonome de la grammaire. Ce type de modèle est notamment illustré par le modèle de Kiparsky (1982). Quand bien même cette composante ne comporte que la morphologie constructionnelle, l'approche de Kiparsky (1982) étant de ce point de vue une approche lexicaliste faible, cette approche accorde tout de même plus d'autonomie à la morphologie que la majorité des autres approches existantes. L'illustration de la figure 2.2 représente les différentes composantes proposées par Kiparsky, le lexique « boîte noire » n'étant plus qu'une sous-partie d'un lexique structuré en termes d'opérations lexicales successives.

Une autre exception, plus notable encore, est la théorie du lexique mise en place par Jackendoff (1975). Cette théorie se situe également dans un cadre lexicaliste faible. Chez Jackendoff, la théorie de l'interface entre morphologie et lexique se résume par ailleurs également à la morphologie constructionnelle. L'approche se distingue néanmoins de la plupart des autres par le fait qu'elle prend effectivement de front la question des



relations morphologiques existantes entre des entrées lexicales distinctes. Elle comporte une modélisation explicite à la fois des entrées lexicales et des règles permettant l'édiction d'entrées nouvelles, appelées RÈGLES DE REDONDANCE (*redundancy rule*) par l'auteur. Chaque entrée est définie par une forme phonologique, une catégorie, des propriétés syntaxiques et des propriétés sémantiques, chacune de ses composantes pouvant être affectée par des règles de redondance. Dans cette thèse, nous présenterons un modèle de la morphologie flexionnelle qui partage avec l'approche de Jackendoff le souci de définition explicite des éléments constitutifs des entrées lexicales. Notre approche se distingue toutefois de la sienne en ce qu'elle formalise non pas les relations entre entrées lexicales mais les propriétés des lexèmes individuels permettant de déduire les caractéristiques de leurs paradigmes — une question qui évidemment ne se pose pas au sein des approches lexicalistes faibles<sup>18</sup>.

Compte tenu de l'assimilation fréquente entre lexique et morphologie, il convient de revenir sur la notion d'autonomie de la morphologie que nous avons évoquée plus haut. La façon dont elle a été définie et défendue par la plupart des auteurs ne stipule réellement qu'une autonomie par rapport à la syntaxe et à la phonologie, pas nécessairement par rapport au lexique.

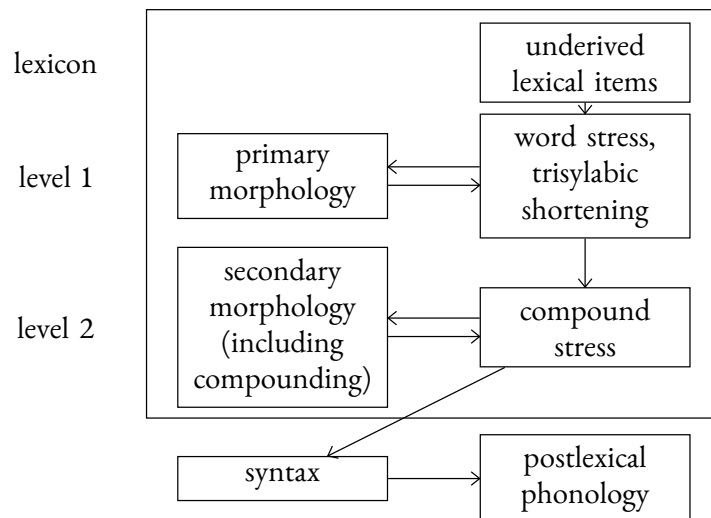


FIGURE 2.2 : Architecture de la grammaire chez Kiparsky (1982)

18. Parmi les autres approches qui pourraient compléter ce rapide panorama des théories du lexique et de son interface avec la morphologie, on pourra également citer les travaux sur les réseaux lexicaux de Bybee (1985) et les travaux sur l'interface entre lexique, morphologie et phonologie de Burzio (2002).

**Théorie du lexique et morphologie autonome** Une théorie n'assigne un statut véritablement autonome à la morphologie que si elle la sépare aussi clairement de la syntaxe que du lexique. Une telle conception permet de faire de la morphologie une sous-partie à part entière de la grammaire, indépendante des autres composantes. Cette séparation entre lexique et morphologie permet également le développement d'une véritable théorie du lexique.

Une approche qui sépare lexique et morphologie est le modèle *Paradigm Function Morphology (PFM)* de Stump (2001, 2006). En PFM, le lexique présente une certaine complexité dans la mesure où il comporte pour chaque lexème un inventaire de radicaux. Ces radicaux sont spécifiés pour être fléchissables par une classe flexionnelle particulière décrite dans une composante morphologique autonome mais pleinement inscrite dans la grammaire. Les classes flexionnelles servent à identifier le champ d'application des règles de réalisations qui permettent la construction des mots forme à partir des radicaux issus du lexique. En revanche, l'approche de Stump reste assez floue sur l'encodage lexical de la notion de lexème par rapport aux radicaux listés dans le lexique. Le rapport entre lexème et radicaux n'est exprimé que dans la théorie annexe de la *Correspondance Paradigmatique (Paradigm Linkage)* décrite par Stump (2006). En dehors de cela, le lexique de Stump ressemble encore partiellement à un lexique bloomfieldien d'exceptions.

Une autre approche qui implicitement sépare la morphologie de la composante lexicale est celle utilisée en morphologie computationnelle dans le formalisme appelé *Morphologie à États Finis (Finite State Morphology)* (Beesley & Karttunen, 2003). Le formalisme repose sur un ensemble d'automates générant les mots. Les modèles stipulent explicitement un lexique de base constitué de morph(èm)es indépendants qui sont ensuite combinés ou modifiés par les règles de réalisation représentées par les automates ou transducteurs appropriés. Cette approche ne comporte donc ni notion de lexème, ni notion de paradigme. La théorie dont la Morphologie à États Finis se rapproche ainsi le plus est sans doute la Morphologie Articulée de Steele (1995). Dans les deux approches, le lexique reste une liste d'éléments non structurée.

*Network Morphology*, théorie lexicaliste forte, lexématique et qui défend explicitement l'autonomie de la morphologie (Brown & Hippisley, 2012, chap. 1.2), comporte une séparation entre lexique et grammaire similaire à celle défendue entre morphologie et syntaxe. *Network Morphology* représente la morphologie comme un ensemble de hiérarchies et de relations de défaut, y compris un héritage par défaut qui relie la morphologie à la syntaxe. Cette dernière consiste entre un héritage orthogonal par défaut entre deux hiérarchies indépendantes ; l'héritage entre lexique et morphologie est du même

ordre (cf. la figure 2.9 page 141). *Network Morphology* définit explicitement une notion de mot WORD qui est hérite de façon orthogonale d'une hiérarchie morphologique pour son comportement morphologique. WORD peut par ailleurs être sous-spécifié selon les différentes catégories de la langue, cf. VERB, NOUN, ADJECTIVE, *etc.* . Chaque mot de la langue hérite ainsi des propriétés morphologiques d'une de ces catégories par défaut<sup>19</sup>. Autrement dit, la hiérarchie qui définit le lexique et la hiérarchie qui définit la morphologie sont deux hiérarchies parallèles distinctes liées par des héritages orthogonaux par défaut.

Le modèle que nous proposons au chapitre 4 comporte une définition explicite de ce qui compose une entrée lexicale et la relie à la composante morphologique, composante autonome de la grammaire. Nous concevons les unités lexicales non pas comme des listes d'éléments imprédictibles dans l'esprit des exceptions bloomfieldiennes, mais comme des éléments complexes porteurs d'information et explicitement interfacés avec les différentes composantes de la grammaire, et en particulier de la morphologie flexionnelle. Nous défendons qu'une théorie attribuant un statut autonome et plein autant au lexique qu'à la morphologie est une théorie d'équilibre dans le partage de l'information.

L'implémentation de notre modèle est réalisée dans une version adaptée du formalisme lexical Alexina développé par Sagot (2005, 2010).

Le formalisme Alexina distingue entre lexique intentionnel et lexique extensionnel. Le lexique intensionnel peut être décrit comme une liste de lexèmes assorties d'informations structurées nécessaires pour déterminer les propriétés morphologiques<sup>20</sup> permettant de construire les formes de leur paradigme. Les règles de réalisation des formes font en revanche partie d'une composante autonome comprenant uniquement la description morphologique (dont en particulier un inventaire des classes flexionnelles pour toutes les catégories, description de toutes les règles de réalisation<sup>21</sup>) de la langue. Nous montrerons dans les chapitres 4 et 6 que pour une même langue la distribution de l'information entre lexique structuré et description morphologique peut varier et que des mesures quantitatives peuvent être appliquées pour optimiser la répartition.

---

19. Lorsqu'une entrée lexicale se distingue du comportement normal de sa catégorie, le cas particulier, spécifié dans la description de l'entrée en question, l'emporte sur le défaut.

20. Dans certains lexiques Alexina comme le *Lefff* pour le français et le *Leffe* pour l'espagnol, le lexique intensionnel comprend également des informations syntaxiques de base, codées dans un format similaire à celui d'un lexique LFG : cadre de sous-catégorisation, redistributions syntaxiques, catégorie syntaxique *etc.* Cette partie ne sera cependant pas utilisée dans les implémentations de notre modèle morphologique.

21. Une nouveauté de la dernière version du formalisme Alexina est la prise en compte de phénomènes non concaténatifs, cf. le chapitre 5 de cette thèse.

## 2.2 Conception modulaire articulée de la grammaire

Dans la section précédente, nous avons discuté les concepts principaux ayant influencé la conception du modèle de la morphologie flexionnelle que nous introduisons dans la seconde partie de cette thèse. Il implémente notamment la notion de morphologie autonome par rapport à la syntaxe en particulier et aux autres composantes de la grammaire en général. Mais il est également explicite quant à l'interaction de la morphologie avec ces autres composantes. Ce fait se manifeste en un premier lieu par la représentation de l'interface entre la morphologie et le lexique. Cependant, notre modèle repose d'une manière générale sur une vision des composantes de la langue comme des modules autonomes les uns des autres tout en étant en constante interaction, ce que nous qualifierons de CONCEPTION MODULAIRE ARTICULÉE DE LA GRAMMAIRE. Nous la représentons de façon simplifiée par la figure 2.3 page 124. L'articulation entre les composantes de la langue y est modélisée par un système de requêtes (lignes rouges) et de réponses (lignes vertes).

Notons que le but du schéma que nous proposons ici est uniquement de rendre plus compréhensible la vision des composantes de la langue qui sous-tend notre modèle. Il ne s'agit en aucun cas du résumé d'une théorie aboutie du fonctionnement de la langue. Nous ne prétendons pas plus qu'il soit satisfaisant pour rendre compte du fonctionnement cognitif de la composante langagière chez l'homme.

Dans le modèle de la grammaire que nous esquissons ici, le lexique comprend pour chacune de ses entrées

- des informations morphologiques (classe flexionnelle, opérations constructionnelles possibles, *etc.* ) ;
- des informations syntaxiques lexicales (catégorie syntaxique, structure argumentale, *etc.* ) ;
- des informations contextuelles (registre, domaine d'utilisation, *etc.* ) ;
- des informations de sémantique lexicale : sens particulier de l'entrée, rapport avec les sens des autres entrées (synonymie, hypéronymie, hyponymie, *etc.* ), appartenance à certaines classes sémantiques (animé/inanimé, comptable/massique, objet, humain, abstrait/concret, *etc.* ) ;
- ...et potentiellement d'autres informations encore.

Notre étude se limitera aux informations morphologiques flexionnelles dans les chapitres à venir.

La syntaxe est responsable des opérations affectant les mots selon les exigences imposées par les composantes sémantique (non lexicale) ou pragmatique, par la structure informationnelle générale ou d'autres données encore liées au contexte de l'énoncé. Les phénomènes de liaison, d'ajustements stylistiques ou autres sont pris en compte par la composante phonologique. Cette dernière peut également répondre à des exigences contextuelles.

La morphologie enfin comporte deux volets distincts que nous avons discutés à la section 2.1.5 : la morphologie constructionnelle, responsable de la construction de lexèmes nouveaux, d'une part, et la morphologie flexionnelle qui réalise les formes fléchies de chaque lexème de l'autre.

La morphologie flexionnelle répond ainsi à la demande de la syntaxe de lui fournir un mot ayant un certain nombre de caractéristiques requises pour la construction de la phrase. Cette requête est la compilation des exigences qu'elle a elle-même reçues des composantes contextuelles. Elles concernent notamment le choix d'une forme d'un lexème avec un sens lexical particulier, adapté à un contexte énonciatif précis et pouvant être utilisé dans une position syntaxique donnée. Ce dernier point permet par exemple de répondre à la requête syntaxique d'un mot pouvant être sujet dans une phrase latine par la réalisation par la morphologie d'une forme nominale au nominatif. Le lexème à fléchir est fourni à la morphologie par le lexique par transmission par la morphologie au lexique de la requête initiale de la syntaxe.

La morphologie constructionnelle, quant à elle, permet la construction de lexèmes nouveaux. À partir de la base d'une des entrées du lexique et sur la requête de la sémantique et du contexte (transmise par la syntaxe et la morphologie), la composante constructionnelle de la morphologie construit ainsi de nouveaux lexèmes qui servent à enrichir le lexique. Ces lexèmes sont ainsi assortis de leur propres informations lexicales (classe flexionnelle, catégorie, sens lexical *etc.*). Ils peuvent à leur tour être fléchis par la composante flexionnelle et leurs formes fléchies peuvent être insérées dans des phrases par la syntaxe. Dans le schéma de la figure 2.3 page 124, nous avons également indiqué une zone TAMPON. Nous considérons en effet que tous les lexèmes créés par la morphologie ne sont pas nécessairement introduits dans le lexique. Certains seront créés suite à des requêtes spécifiques (avec, entre autre, leur information morphologique flexionnelle), fléchis à la volée et insérés en syntaxe. Mais leur caractère exceptionnel ne justifie pas encore à ce stade leur introduction dans le lexique en tant que nouvelle entrée lexicale disponible pour des requêtes ultérieures. Cela ne sera établi que par un processus de lexicalisation spécifique (flèche bleue) pour les lexèmes dont la fréquence d'emploi et/ou la difficulté à les dériver de

façon régulière dans la composante constructionnelle le justifient — en termes d'économie du système global selon notre critère. Dans la suite de cette thèse, nous n'approfondirons cependant pas la dimension constructionnelle de la morphologie. Nous laissons donc les questions d'arbitrage en termes de lexicalisation pour des études ultérieures sur le sujet.

La figure 2.3 résume ainsi l'esquisse de l'architecture de la grammaire que nous avons proposée et sur laquelle repose la vision de la morphologie que nous développons pour sa composante flexionnelle dans la suite de cette thèse.

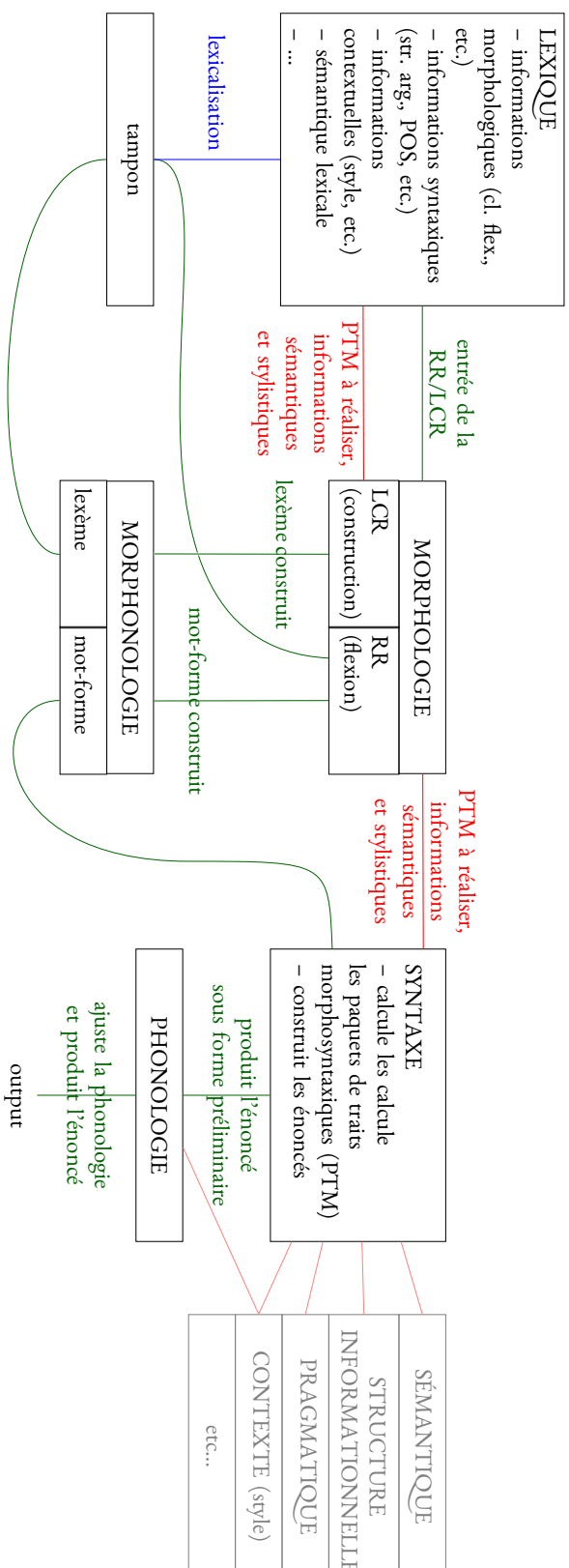


FIGURE 2.3 : Esquisse de l'architecture de la grammaire

## 2.3 Approches inférentielles-réalisationnelles

Dans la dernière section de ce chapitre, nous présentons rapidement les caractéristiques majeures des deux modèles qui ont le plus directement influencé le modèle que nous introduisons au chapitre 4 : *Paradigm Function Morphology* (Stump, 2001, 2006) et *Network Morphology* (Corbett & Fraser, 1993 ; Brown & Hippiusley, 2012). Ces deux modèles partagent avec le nôtre un certain nombre de caractéristiques, dont le fait d'appartenir aux approches lexématiques et de postuler l'autonomie de la morphologie.

Par ailleurs, ces modèles adoptent tous les trois une approche inférentielle-réalisationnelle de la morphologie<sup>22</sup> selon la typologie des approches formelles de la morphologie flexionnelle proposée par Stump (2001) en introduction à la présentation de son propre formalisme. Dans cette typologie, les approches formelles de la morphologie flexionnelle se distinguent selon deux axes : 1) une approche peut être soit *lexicale* soit *inférentielle* ; 2) elle peut être *incrémentale* ou *réalisationnelle*, cf. les définitions ci-dessous.

**Lexical** Les marques flexionnelles sont des éléments du lexique au même titre que tout autre élément (racines, radicaux, thèmes *etc.* .)

**Inférentiel** Les marques flexionnelles sont l'expression de traits morphosyntaxiques par des règles de réalisation spécifiques apparant un mot et sa racine.

**Incrémental** La morphologie flexionnelle augmente le contenu informationnel d'un mot, c'est-à-dire que l'acquisition de propriétés morphosyntaxiques est concomitant à l'acquisition d'exposants flexionnels.

**Réalisationnel** L'association d'un mot avec des propriétés morphosyntaxiques particulières conduit à l'introduction des exposants flexionnels de ces propriétés.

Le tableau 2.6 indique des exemples de modèles que nous avons évoqués dans les sections précédentes et qui correspondent chacun à un des quatre types d'approches définies par Stump. Notons que les approches Mot et Paradigme se situent toutes dans la case inférentielle-réalisationnelle.

Dans les présentations que nous faisons de PFM et *Network Morphology* ci-dessous, nous nous intéressons par ailleurs en particulier aux points suivants qui seront centraux dans la conception de notre modèle :

1. Comment et avec quelle force se manifeste le caractère lexématique des approches sous-jacentes aux modèles ?

---

22. Ce point n'étant cependant pas la question centrale dans le modèle que nous introduisons au chapitre 4.



	LEXICAL	INFÉRENTIEL
INCRÉMENTAL	(Lieber, 1992)	<i>Articulated Morphology</i> (Steele, 1995)
RÉALISATIONNEL	<i>Distributed Morphology</i> (Halle & Marantz, 1993)	<i>Extended Word &amp; Paradigm</i> (Anderson, 1982); <i>Paradigm Function Morphology</i> (Stump, 2001); <i>Network Morphology</i> (Brown & Hippisley, 2012)

TABLE 2.6 : Typologie des approches formelles de la morphologie (Stump, 2001)

2. Comment l'autonomie de la morphologie apparaît-elle dans la modélisation ?
3. Quel est le rapport entre le lexique et la grammaire ? Quels sont les éléments reliant les deux ?
4. Comment le caractère inférentiel-réalisationnel est-il exprimé ?
5. Quel est le statut des traits morphosyntaxiques dans le modèle et où et comment interviennent-ils ?
6. Quelle est la valeur de la notion de défaut dans les modèles ?
7. Comment des règles spécifiques, ou *exceptions*, peuvent-elles se substituer aux défauts ?
8. Quelles sont les implémentations existant pour ces modèles ?

### 2.3.1 *Paradigm Function Morphology* Stump (2001)

*Paradigm Function Morphology* ou PFM est un formalisme introduit pour la première fois par Stump (1991). Le livre de référence à ce jour est *Inflectional Morphology : A Theory of Paradigm Structure* de Gregory T. Stump (2001). PFM peut être conçu comme une élaboration et une formalisation d'*Extended Word & Paradigm (EWP)* (Mot et Paradigme Étendu) développé par (Anderson, 1982).

PFM est une théorie de la morphologie flexionnelle dans un cadre Mot et Paradigme. S'il s'agit en un premier lieu d'un formalisme caractérisant spécifiquement la forme des règles de réalisation, il ne faut pas pour autant confondre le formalisme mis en place pour rendre compte des règles de réalisation et la théorie sous-jacente qui a une portée bien plus large que la simple réalisation formelle des mots.

### 2.3.1.1 Le formalisme associé à PFM

Le formalisme associé à PFM est un formalisme réalisationnel qui emploie la notion de *structure de traits*. Une structure de traits est un ensemble de couples de traits, c'est à dire de couples attribut-valeur ; par exemple *pluriel* et *singulier* sont des valeurs possibles de l'attribut NOMBRE<sup>23</sup>.

Le fonctionnement d'une règle de réalisation en PFM repose sur le mécanisme de l'unification de structures de traits. Les règles réalisées sont celles compatibles avec les paquets de traits morphosyntaxiques pour lesquelles une case est spécifiée.

Il existe deux sortes de règles de réalisation : les règles d'exponence et les règles de renvoi. On note une règle de réalisation de type règle d'exponence comme en 2.15<sup>24</sup>.

$$2.15 \quad X_C, \sigma \longrightarrow f(X)$$

La règle en 2.15 comporte ainsi les indications suivantes :

1. Une forme de départ notée  $X$  ;
2. La classe de lexème  $C$  à laquelle appartient  $X$  et notée en indice de  $X$ , cf.  $X_C$  ;
3. L'indication des traits réalisés par la règle, notés  $\sigma$  et habituellement notés sous la forme de structures de traits ;
4. Une fonction qui calcule la sortie (ou forme d'arrivée) de la règle, notée  $f(X)$ .

Une règle réalise ainsi pour un lexème donné appartenant à une classe de lexèmes  $C$  les traits morphosyntaxiques indiqués par  $\sigma$  en appliquant à une forme de départ  $X$  la transformation formelle définie par la fonction  $f(X)$ . Ces transformations peuvent aller de la simple affixation jusqu'à n'importe quel autre changement plus complexe sur la séquence phonologique<sup>25</sup> en passant par les phénomènes d'*apophonie* (*Ablaut*<sup>26</sup> et *Umlaut*<sup>27</sup>) et de *transfixation*<sup>28</sup> jusqu'à la *réduplication* (complète ou partielle<sup>29</sup>).

23. Par convention, on note souvent le nom des attributs en petites majuscules et le nom des valeurs en minuscules. Un trait, couple attribut-valeur, sera ainsi noté {ATTRIBUT *valeur*}.

24. La notation adoptée ici est celle de Ackerman & Stump (2004). Pour une synopsis à jour de PFM, mais qui adopte une notation légèrement différente, voir Bonami & Stump (à paraître).

25. Ou graphémique si l'on veut traiter de l'écrit.

26. Cf. Le changement vocalique des verbes forts en anglais comme SING « chanter » qui donne *sing*, *sang* et *sung* pour le présent, le passé et le participe respectivement.

27. Par exemple les changements vocaliques dans la construction des pluriels de certains noms en allemand, cf. pour BRUDER « frère » qui donne *Bruder* au singulier et *Brüder* au pluriel. La différence avec l'Ablaut est que l'Umlaut correspond à un changement phonétique régulier, en l'occurrence la transformation d'une voyelle postérieure en voyelle antérieure.

28. Comme pour la flexion de l'arabe, cf. McCarthy (1982).

29. Cf. la définition de Rubino (2011) citée à la section 3.1.1 du chapitre 3.

L'application des règles est évaluée en fonction de la classe  $c$  de lexèmes à laquelle appartient le lexème de la base  $X$  en question, ainsi qu'en fonction de l'adéquation entre les traits morphosyntaxiques exprimés par le mot réalisé et les traits morphosyntaxiques  $\sigma$  réalisés par la règle. Une règle peut ainsi s'appliquer à une base si la structure de traits  $\sigma$  indiquant les traits réalisés par la règle subsume la structure de traits réalisée par le mot ; ou autrement dit, si la structure de traits correspondant aux traits réalisés par le mot est une extension de  $\sigma$ , la structure de traits représentant les traits réalisés par la règle.

Dans les cas où un mot porte plusieurs marques flexionnelles introduites par des règles d'exponence distinctes, plusieurs règles sont appliquées de façon séquentielle. La sortie de la règle précédente devient ainsi la forme de base de la suivante<sup>30</sup>.

Comme EWP (Anderson, 1982), PFM repose par ailleurs sur un système de blocs de règles. Ces blocs ordonnés contenant les règles de réalisation doivent s'appliquer dans l'ordre : le bloc 1 avant le bloc 2 et ainsi de suite<sup>31</sup>.

Les blocs permettent de rendre compte des systèmes de *classes de position* qui représentent des exposants mutuellement exclusifs. Nous donnons un exemple de classes de positions pour le système verbal du kurde sorani<sup>32</sup>. Les paradigmes flexionnels des verbes sorani suggèrent la présence de sept classes de position. Les sept classes de position sont répertoriées dans le tableau 2.7<sup>33</sup> Dans une formalisation PFM, une classe de position comprenant des préfixes et une classe comprenant des suffixes peuvent néanmoins être réalisées dans un seul bloc, à condition que les deux types de marques ne soient mutuellement exclusifs. Ces sept classes peuvent être prises en charge par cinq blocs de règles<sup>34</sup>.

Dans chaque bloc une et une seule règle s'applique à une forme de base donnée. La règle correspondante est sélectionnée en fonction de l'adéquation de la classe de lexème indiquée dans la règle et en fonction de l'adéquation à la structure de traits spécifiée par la règle.

Une caractéristique marquante de PFM est qu'elle fait également usage d'une notion de compétition entre règles. Dans un bloc donné, si plusieurs règles sont compatibles

---

30. Notons qu'une implémentation des règles PFM par des automates à états finis est possible pour tous les phénomènes hormis la reduplication, le phénomène de copie ne pouvant être représenté par des automates sans les étendre par des mécanismes complémentaires (Malouf, 2005). Notons encore qu'une telle implémentation pourrait alors tout simplement être réalisée par un seul et même automate obtenu par la composition des automates individuels correspondant aux règles particulières dans l'ordre adapté.

31. Voir aussi la Morphologie Articulée de Steele (1995).

32. Langue indo-iranienne de l'ouest parlée essentiellement à la frontière irano-irakienne, mais aussi au nord de la Syrie.

33. Voir la section 6.4 au chapitre 6 pour plus de détails sur les désinences indiquées dans le tableau.

34. Par exemple : bloc 1 pour les classes -1 et 1, bloc 2 pour les classes -2 et 2, bloc 3 pour la classe 3, bloc 4 pour la classe 4 et bloc 5 pour la classe 5.

-2	-1	0	1	2	3	4
mə	də	RADICAL (1, 2, 3, 4)	ra	u(:)	m	ayə
nə	b		re	bu:	i(t)	
na				b	e(t)	
					ə	
					in	
				n		
			n			

TABLE 2.7 : Classes de position des verbes sorani

avec une entrée donnée (classe de lexèmes et structure de traits compatibles), la règle la plus spécifique l'emporte sur les autres. Le principe permettant de choisir la règle la plus spécifique dans un bloc donné est appelé principe de Paṇini (Stump, 2001, p. 23). L'évaluation de la règle la plus spécifique se fait formellement sur les ensembles de traits concernés par la règle et la classe de lexèmes concernés par la règle. Le fait pour une règle d'être applicable à une entrée en termes de compatibilité de structures de traits se traduit par le fait que la structure de traits indiquant les traits réalisés par la règle doit subsumer la structure de traits réalisée par le mot. Dans la compétition entre règles, ce sera la règle qui, parmi les règles subsumant la structure de traits du mot à réaliser, sera l'extension maximale de toutes les autres qui s'appliquera. Dans une théorie PFM bien formée, il est toujours possible de déterminer sans ambiguïté la règle la plus spécifique d'un bloc.

De ce point de vue PFM se distingue d'EWP (Anderson, 1982). En EWP, les règles sont ordonnées à l'intérieur des blocs de règles. À l'intérieur d'un bloc donné, ce sera alors toujours la première règle compatible avec les traits à exprimer par une forme qui sera appliquée, sans qu'une comparaison inter-règles ne soit nécessaire. En revanche, il est dès lors nécessaire, en EWP, de spécifier un ordre de préséance externe entre les règles d'un même bloc afin d'éviter que la mauvaise règle ne soit réalisée pour une forme donnée. D'un point de vue formel, les deux systèmes permettent néanmoins d'obtenir les mêmes résultats : PFM nécessite une spécification suffisante des traits exprimés par les règles pour garantir le succès du principe de Paṇini ; EWP nécessite un ordonnancement externe des règles.

PFM stipule également des règles dites *porte-manteau*. Elles ont la particularité de réaliser en une seule règle un exposant correspondant à deux classes de positions distinctes, cf. le préfixe négatif na dans le tableau 2.7. Les règles porte-manteau sont prioritaires sur les règles simples. Stump (2001, p. 142) introduit la règle dite règle de *Composition*

*Fonctionnelle par Défaut* (*Function Composition Default (FCD)*) dont la fonction est de mettre les combinaisons de règles simples en compétition avec les règles porte-manteau.

$$2.16 \quad f[n,m](\langle X, \sigma \rangle) =_{def} f_n(f_m(\langle X, \sigma \rangle))$$

Enfin, PFM comporte également un mécanisme permettant de rendre compte des syncrétismes directionnels à l'intérieur d'un paradigme. Un cas de syncrétisme standard est le syncrétisme entre l'accusatif et le nominatif au neutre en latin, cf. le tableau 2.8. Le tableau 2.8 présente pour comparaison les nominatifs et accusatifs singulier et pluriel de trois noms latins : FILIUS « *fil*s » (masculin), FILIA « *fil*le » (féminin), et DONUM « *don* » (neutre). Seuls les noms neutres présentent un syncrétisme entre nominatif et accusatif autant au singulier qu'au pluriel.

		FILIUS	FILIA	DONUM
		M	F	N
S	NOM	<i>filius</i>	<i>filia</i>	<i>donum</i>
	ACC	<i>filium</i>	<i>filiam</i>	<i>donum</i>
P	NOM	<i>fili</i>	<i>filia</i>	<i>dona</i>
	ACC	<i>filios</i>	<i>filias</i>	<i>dona</i>

TABLE 2.8 : Nominatifs et accusatifs de trois noms latins

En PFM, les syncrétismes sont formalisés par des *règles de renvoi* (Zwicky, 1985). Une règle de renvoi permet d'indiquer pour un ensemble de traits que la forme correspondante est réalisée comme celle qui est associée à un autre ensemble de traits.

En PFM, les règles de renvoi sont notées comme en 2.17.

$$2.17 \quad X_N, \sigma : \{\text{GENRE } n, \text{CAS } acc\} \longrightarrow \langle X, \sigma / \{\text{CAS } nom\} \rangle$$

Une telle règle stipule que la règle de réalisation associée au nominatif pour les noms (classe N) neutres en latin est la même que celle associée à l'accusatif dans le bloc 1. En particulier, cette règle indique que la règle permettant de produire la bonne forme fléchie est appliquée pour un ensemble de traits qui ne diffère des traits de la première que pour le trait spécifié, à savoir ici {CAS *acc*} au lieu de {CAS *nom*}.

Les règles de renvoi peuvent aussi bien agir à l'intérieur d'un même bloc qu'entre blocs. Les règles qui renvoient à la réalisation à l'intérieur d'un bloc distinct sont appelées *règles*

*trans-bloc*. Elles spécifient, en plus des traits d'entrée modifiés, le bloc auquel appartient la règle permettant de réaliser la forme correcte.

Les mécanismes que nous avons présentés dans cette sous-section constituent l'appareil formel que comporte *Paradigm Function Morphology*. Mais PFM ne se réduit pas à ces formalisations. PFM comporte par ailleurs une théorie étendue du domaine de la morphologie flexionnelle.

### 2.3.2 La théorie flexionnelle en PFM

Au-delà du fait de comporter un formalisme explicite de la réalisation des formes fléchies, PFM est également une théorie linguistique de la morphologie flexionnelle.

**La théorie originelle de Stump (2001)** En tant que théorie spécifique à la morphologie flexionnelle, PFM comporte une définition explicite de la notion de flexion : Dans une phrase, chaque mot porte un ensemble de traits morphosyntaxiques appelé ENSEMBLE DE PROPRIÉTÉS MORPHOSYNTAXIQUES (Stump, 2001, p. 1) ; dans une très grande partie des cas, ces propriétés se reflètent dans la morphologie du mot en question.

Pour Stump la forme d'un mot est la manifestation matérielle de ces propriétés morphosyntaxiques. Elle est obtenue par l'application de règles de réalisation appliquées à un radical d'un lexème donné pour générer la forme concrète de son paradigme qui correspond à ces propriétés<sup>35</sup>.

PFM s'accompagne notamment d'une théorie de la notion de lexique. Le lexique flexionnel comporte ainsi une collection de radicaux<sup>36</sup>. Chacun de ces radicaux est associé à une catégorie morphosyntaxique qui peut elle-même comporter plusieurs sous-catégories. Modéliser en PFM les classes flexionnelles verbales du latin décrites au chapitre 1, reviendrait ainsi à dire que les quatre classes flexionnelles que nous avons définies sont associées à la catégorie verbe et que chaque catégorie comporte un certain nombre de radicaux dont elle définit la flexion.

PFM permet par ailleurs de regrouper certaines de ces classes qui partagent des caractéristiques communes sous l'étiquette d'une super-classe, ce qui présente l'avantage de simplifier l'écriture des règles, une seule règle étant ainsi compatible avec deux classes. L'illustration de cette représentation est donnée à la figure 2.4 page 132 et dans les règles PFM indiquées dans l'annexe A.

---

35. Cf. la définition des approches inférentielles-réalisationnelles ci-dessus p. 127.

36. Appelés RACINES dans (Stump, 2001).

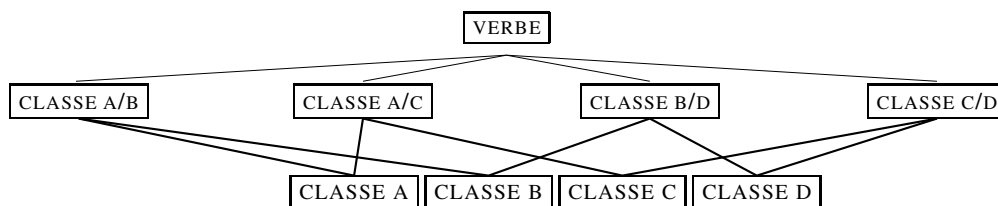


FIGURE 2.4 : Représentation des super-classes en PFM

Nous avons montré que PFM (Stump, 2001) permet l'expression explicite<sup>37</sup> des règles de réalisations construisant les formes du paradigme d'un lexème donné. Le modèle permet par ailleurs la représentation formelle des phénomènes de syncrétismes. Ces égalités entre formes d'un même paradigme sont exprimées par des règles de renvoi. Le but de ce mécanisme n'est pas le simple encodage des formes ; les règles simples seraient suffisantes pour cela. Par ce mécanisme, PFM cherche à encoder explicitement des caractéristiques formelles non seulement des formes individuelles, mais aussi du paradigme dans son ensemble. PFM permet ainsi d'exprimer formellement des caractéristiques paradigmatiques à proprement parler.

La correspondance paradigmatique Stump (2006) associe PFM à la théorie dite PARADIGM LINKAGE (PL) (*Correspondance paradigmatique*). La théorie flexionnelle décrite dans (Stump, 2006) va au delà de ce qui est dit dans (Stump, 2001). Cette extension permet notamment de rendre compte des phénomènes de décalage entre formes fléchies et propriétés morphosyntaxiques exprimées au sens de Baerman (2007b). Stump représente ainsi un lexème comme un couple de deux paradigmes, l'un étant le *paradigme de contenu* (*Content Paradigm*), l'autre le *paradigme de forme* (*Form Paradigm*).

Le paradigme de contenu contient les formes requises par la syntaxe, à savoir des couples opaques (lexème, propriétés morphosyntaxiques exprimées) notés  $\langle L, \sigma \rangle$ . Le paradigme de forme correspond à ce qui est réalisé par les règles d'exponence. Il contient des couples (radical, propriétés morphosyntaxiques réalisées) notés  $\langle s, \tau \rangle$ . La combinaison des informations contenues dans ces deux paradigmes définit les propriétés des mots-formes, cf. l'illustration de la figure 2.5.

Par défaut les propriétés morphosyntaxiques exprimées par le paradigme de contenu sont les mêmes que celles réalisées par le paradigme de forme. Cette correspondance correspond à ce que Stump (2006) appelle la RÈGLE DE CORRESPONDANCE PARADIGMATIQUE

37. Voir la description du formalisme de PFM ci-dessus.

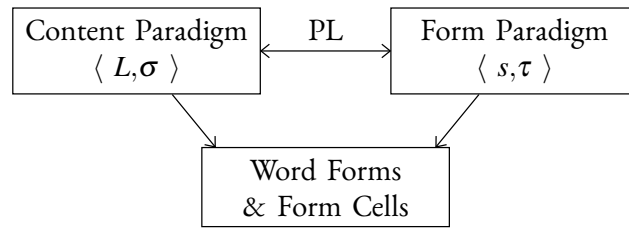


FIGURE 2.5 : Constructions des mots et *Correspondance Paradigmatique*

PAR DÉFAUT (*Default Rule of Paradigm Linkage*).

Si jamais les propriétés morphosyntaxiques réalisées par un couple  $\langle s, \tau \rangle$  du paradigme de forme ne correspond pas à celles réalisées par le couple  $\langle L, \sigma \rangle$  du paradigme correspondant, nous obtenons une déviation par rapport à la règle de correspondance paradigmatique par défaut. Cette déviation par rapport à la règle a pour résultat l'apparition de phénomènes de décalage entre forme et fonction au sens de Baerman (2007b)<sup>38</sup>.

Stump (2006) analyse ainsi le comportement des verbes déponents en latin. Adoptant, comme Hippisley (2007), une approche flexionnelle du passif, il considère que l'utilisation de formes morphologiques passives par les verbes déponents pour exprimer une valeur syntaxiquement active correspond à un décalage entre forme et fonction. Ce décalage est alors modélisé par une règle de correspondance paradigmatique spécifique illustrée ci dessous.

2.18 (Stump, 2006, p. 288)

Définition de mise en correspondance des propriétés en latin :

Si  $\sigma = \{\text{ACTIF X}\}$ , alors  $f(\sigma) = \{\text{PASSIF X}\}$ ;

sinon  $f(\sigma) = \sigma$ .

Les caractéristiques principales de PFM se résument par les points suivants.

1. Structure des règles de réalisation :

- a) Les mots sont construits par des règles de réalisation
- b) Les règles de réalisation sont organisées par blocs de règles.
- c) À l'intérieur d'un bloc, les règles sont en compétition les unes avec les autres. Parmi les règles compatibles avec les traits d'un mot-forme à réaliser, la règle la plus spécifique l'emporte sur les autres.

38. Voir aussi la définition de la notion de *décalage* au chapitre 3.



- d) Des règles de renvoi permettent d'encoder explicitement les phénomènes de syncrétisme à l'intérieur des paradigmes.
2. Rapport au lexique (Stump, 2001) :
- a) Le lexique de PFM comporte des radicaux.
  - b) Le lexique est structuré en classes de mots qui reposent sur les catégories morphosyntaxiques.
  - c) Il comporte également des super-classes permettant d'exprimer des informations concernant plusieurs sous-classes à la fois.
  - d) Extension du rapport au lexique (Stump, 2006) :
  - e) Stump (2006) en plus différencie entre un paradigme de contenu et un paradigme de forme.
  - f) Le paradigme de contenu correspond aux informations lexématiques. Le paradigme de forme est propre à un radical unique.
  - g) Une règle de Correspondance Paradigmatique permet de relier les deux paradigmes, et de ce fait le lexème à son radical (ou ses radicaux multiples).

### 2.3.2.1 Caractéristiques particulières de PFM

**Morphologie lexématique** En tant que théorie appartenant aux approches Mot et Paradigme, on peut considérer PFM comme une théorie lexématique. Toutefois, il est à noter que dans (Stump, 2001) les entrées lexicales tout comme les éléments donnés en entrée aux règles de réalisation ne sont pas les lexèmes en tant que tels mais un des radicaux du lexème en question.

**Autonomie de la morphologie** « L'exponence est la seule association entre les marques flexionnelles et les propriétés morphosyntaxiques » Stump (2001, p. 11). Les propriétés morphosyntaxiques sont utilisées en syntaxe pour exprimer les phénomènes d'accord ou de gouvernement, mais la syntaxe n'a accès qu'aux mots entièrement fléchis assortis de leurs propriétés morphosyntaxiques.

**Lexique et grammaire** Dans (Stump, 2001), le lexique correspond à un inventaire de radicaux. Dans (Stump, 2006), le lien entre lexique et grammaire en PFM s'opère par la correspondance entre les paradigmes de contenu et les paradigmes de forme. Cette correspondance permet de représenter des caractéristiques morphologiques comme les

décalages ou l'hétéroclise (cf. chapitres 3 et 4) par des relations de correspondance lexicales entre un lexème donné et les radicaux stockés dans le lexique.

Par ailleurs, les héritages entre catégories morphosyntaxiques, classes flexionnelles et super-classes permettent d'exploiter dans les règles de réalisation de la composante morphologique de la grammaire des relations hiérarchiques entre unités lexicales établies dans le lexique.

**Théorie inférentielle-réalisationnelle** Les règles d'exponence réalisent seules l'association observable entre les propriétés morphosyntaxiques portées par un mot dans un énoncé et les marques flexionnelles identifiables dans la forme du mot.

**Statut des traits morphosyntaxiques** Les traits morphosyntaxiques sont en général représentés sous forme de structures de traits dans la modélisation associée à PFM. Elles déterminent l'applicabilité des règles de réalisation et servent à résoudre la compétition entre règles.

Les traits morphosyntaxiques établissent également le lien entre le paradigme de contenu et le paradigme de forme.

**La notion de défaut** La notion de défaut est fondamentale dans le modèle théorique associé à PFM. Elle apparaît autant dans la conception de la théorie que dans l'architecture de l'appareil formel.

Dans la théorie de correspondance paradigmatique, le *Défaut de Correspondance Paradigmatique par Défaut*, Default Rule of Paradigm Linkage, est une règle qui relie, par défaut, entre eux le paradigme de contenu et le paradigme de forme d'un lexème. Toute déviation par rapport à cette règle doit être spécifiée par une fonction de mise en correspondance du paradigme de contenu avec le paradigme de forme.

PFM comporte par ailleurs deux mécanismes permettant d'introduire des défauts dans l'appareil formel :

La *Fonction d'Identité par Défaut*, Identity Function Default (IFD), stipule que si aucune règle de réalisation n'est stipulée pour une certaine combinaison de traits particulière dans la description, la sortie de la réalisation de la forme concernée est identique à l'entrée. Cette règle permet ainsi d'alléger la description.

La règle de *Composition Fonctionnelle par Défaut*, Function Composition Default (FCD), stipule qu'une règle porte-manteau correspond à la composition de deux (ou plus de deux) règles appartenant à autant de blocs et est par définition plus spécifique que les règles simples ne correspondant qu'à un seul bloc.

**Traitement des exceptions** Les défauts mis en place par les mécanismes ci-dessus peuvent être annulés par des spécifications supplémentaires. La règle IDF est annulée à chaque fois qu'une autre règle, explicite, est compatible avec les traits exprimés par un mot donné. Ce cas correspond ainsi à un cas particulier du principe de Paṇini régissant la compétition entre règles et selon lequel la plus spécifique l'emporte sur les autres. L'IDF est par définition minimalement spécifiée : elle correspond à la règle associée à une structure de traits vides, subsumant par définition toutes les autres <sup>39</sup>.

**Modes d'implémentation** En collaboration avec Gregory T. Stump, Raphael Finkel a développé *Cat's Claw*, un système d'implémentation spécialement adapté à PFM <sup>40</sup>. Ce système comporte un ensemble d'outils, dont KATR, une extension de DATR adaptée à la théorie de PFM, et un fléchisseur utilisant des analyses PFM explicites et un analyseur en parties principales. Quand bien même ces outils sont en libre accès sur le site du projet, le code des outils supplémentaires ne semble malheureusement pas être disponible en ligne.

### 2.3.3 *Network Morphology* Brown & Hippiisley (2012)

#### 2.3.3.1 Présentation générale de *Network Morphology*

L'origine de *NETWORK MORPHOLOGY* remonte aux travaux de Corbett & Fraser (1993). La description la plus complète du modèle est faite par Dunstan Brown et Andrew Hippiisley dans le livre intitulé *Network Morphology : A Defaults-based Theory of Word Structure* (Brown & Hippiisley, 2012).

Contrairement à PFM, *Network Morphology* n'est pas un modèle de la seule morphologie flexionnelle. Parmi les études importantes effectuées en *Network Morphology* figure la thèse d'Andrew Hippiisley (1997) sur la dérivation en russe.

Par rapport à PFM, *Network Morphology* semble aussi bien plus intimement liée au formalisme qui lui sert d'implémentation. Une théorie *Network Morphology* s'écrit toujours en DATR (Evans & Gazdar, 1989, 1996). Cette proximité est d'ailleurs parfaitement assumée par les auteurs qui considèrent l'emploi de techniques issues de la morphologie computationnelle comme moyen de vérification de la validité de théories morphologiques <sup>41</sup>.

---

39. On pourrait par ailleurs arguer que le fait qu'une règle PFM puisse spécifier une sous-classe de lexèmes à laquelle elle s'applique constitue une exception au défaut qu'une règle s'applique si elle réalise les traits morphosyntaxiques portés par le mot.

40. Les outils *Cat's Claw* sont disponibles sur la page du projet (*Computer-Assisted Service : Computational Linguist's Automated Workbench*) à l'adresse [http://www.rch.uky.edu/project\\_claw.html](http://www.rch.uky.edu/project_claw.html) (valable au 11 juin 2013).

41. Voir notamment les pages du projet *Network Morphology* et du livre (Brown & Hippiis-

*Network Morphology* repose sur l'idée que la langue, et en particulier la morphologie et le lexique, peut être représentée sous la forme de hiérarchies ou réseaux d'héritage. La grammaire complète est conçue comme une cascade de défauts qui dans certains cas particuliers peuvent être remplacés par des exceptions.

Par exemple, si tous les noms d'une langue partagent le fait de se fléchir en nombre, un nom particulier héritera par défaut de cette caractéristique. Par exemple, en français, les noms se fléchissent par défaut en nombre et peuvent exprimer deux valeurs pour l'attribut NOMBRE, à savoir le *singulier* ou le *pluriel*. Un nom comme VIVRES qui ne présente qu'une forme plurielle constitue de ce point de vue une exception qui doit être marquée par le blocage de l'héritage par défaut dans la hiérarchie flexionnelle du français.

Enfin, *Network Morphology* est le formalisme principal dans lequel ont été traités formellement les travaux appartenant aux approches canoniques (Corbett, 2003) de la typologie, et en particulier de la morphologie.

### 2.3.3.2 Caractéristiques particulières de *Network Morphology*

**Morphologie lexématique** Brown & Hippisley (2012, p. 10-20) défendent explicitement le fait d'adopter une approche lexématique plutôt que morphématique. Ce choix est une conséquence directe du fait que la théorie adopte une approche Mot et Paradigme et fait du domaine du paradigme l'un des domaines dans lesquels peuvent être stipulées des généralisations sur la langue étudiée.

Ainsi, *Network Morphology* définit explicitement dans son lexique des entrées lexicales qui ont le statut de lexème. Les lexèmes sont organisés dans une hiérarchie en fonction des catégories morphosyntaxiques auxquelles ils appartiennent. Chaque lexème appartenant à une catégorie morphosyntaxique donnée hérite ainsi (par défaut et à moins d'être soumis à une exception, cf. plus bas) de l'intégralité des informations spécifiées pour cette catégorie.

Chacun de ces lexèmes est défini par un nœud dans la théorie DATR. Ce nœud peut ainsi définir les informations suivantes pour chaque lexème.

1. le nœud supérieur dont hérite l'entrée : il s'agit souvent de la catégorie morphosyntaxique de l'entrée en question, ou, à défaut, d'un sous-ensemble de cette catégorie se fléchissant selon une classe flexionnelle particulière ;
2. la racine de l'entrée en question ;
3. une glose qui est souvent une traduction ;

---

ley, 2012), [http://www.rch.uky.edu/project\\_networkmorphology.html](http://www.rch.uky.edu/project_networkmorphology.html) et <http://networkmorphology.as.uky.edu> : « *The Network Morphology philosophy promotes the use of computational implementation to check theories* ». (valable au 11 juin 2013)

4. un radical au sens de Matthews (1974), c'est-à-dire la racine augmentée d'un thème particulier ;
5. la catégorie syntaxique de l'entrée en question – en particulier lorsqu'elle est différente de la catégorie morphologique de l'entrée (cf. les participes employés comme adjectifs en français) ;
6. une catégorie sémantique contenant des traits de type [ $\pm$  animé], [ $\pm$  animé], *etc.* ;
7. ...et éventuellement d'autres informations encore.

Parmi ces informations, seule celle qui définit le nœud supérieur est obligatoire. La racine et/ou le radical sont également en général spécifiés. Nous donnons deux exemples d'entrées lexicales en *Network Morphology* dans les figures 2.6 et 2.7. Les exemples sont des entrées proposées par Brown & Hippisley (2012) : respectivement l'entrée RUKOP'IS' du russe<sup>42</sup> et AMO du latin<sup>43</sup>.

Dans ces entrées, le nom RUKOP'IS' n'hérite pas directement de sa catégorie mais d'un sous-ensemble des noms qui se décline selon la classe flexionnelle N\_III. L'entrée comporte ici par ailleurs des indications sur sa catégorie syntaxique (<syn cat> == n), son sens par l'intermédiaire de la glose (<gloss> == manuscript) et son radical (<stem> == rukop'is').

```

1 Rukop'is' :
2   <> == N_III
3   <syn cat> == n
4   <gloss> == manuscript
5   <stem> == rukop'is' .

```

FIGURE 2.6 : Entrée lexicale du russe en *Network Morphology*

L'entrée AMO hérite directement de sa catégorie (VERB) et comporte par ailleurs des indications sur son sens par l'intermédiaire de la glose (<gloss> == love). L'entrée indique également la racine associée au lexème (<root> == am) et un nœud supérieur dont l'entrée hérite les informations nécessaires à la construction de son radical (<stem> == CONJ\_1), en l'occurrence le fait d'appartenir aux groupe de verbes à voyelle thématique -a- codé comme CONJ\_1.

*Network Morphology* est une théorie explicitement lexématique.

42. [http://networkmorphology.as.uky.edu/sites/default/files/ch2\\_polishlex\\_0.dtr](http://networkmorphology.as.uky.edu/sites/default/files/ch2_polishlex_0.dtr) (valable au 11 juin 2013).

43. [http://networkmorphology.as.uky.edu/sites/default/files/ch5\\_latin.dtr](http://networkmorphology.as.uky.edu/sites/default/files/ch5_latin.dtr) (valable au 11 juin 2013).

```

1 Amo:
2   <> == VERB
3   <gloss> == love
4   <root> == am
5   <stem> == CONJ_1.

```

FIGURE 2.7 : Entrée lexicale du latin en *Network Morphology*

**Autonomie de la morphologie** L'autonomie de la morphologie est explicitement défendue par Brown & Hippisley (2012, p. 20-47). Comme dans le cas de PFM, il est considéré en *Network Morphology* que les éléments traités en syntaxe sont des mots entièrement fléchis. Les sorties des règles morphologiques sont entièrement opaques pour la syntaxe.

*Network Morphology* considère néanmoins qu'il existe des relations de défaut entre morphologie et syntaxe. Par exemple, un mot appartenant morphologiquement à la catégorie des noms appartient par défaut à la classe des noms en syntaxe.

**Lexique et grammaire** Les analyses *Network Morphology* comportent en général des hiérarchies distinctes mais reliées pour représenter le lexique d'une part et la grammaire, en l'occurrence la morphologie, de l'autre.

**Théorie inférentielle-réalisationnelle** Les formes sont les réalisations de traits morphosyntaxiques définissant les règles de réalisations appartenant à une classe flexionnelle donnée. Au moment de la génération de la théorie DATR associée à la description morphologique, les formes produites par la théorie sont celles qui ont été appelées par spécification des traits morphosyntaxiques pertinents. Cet appel est représenté par la commande `#show` en DATR. Nous en donnons un exemple dans la figure 2.8 extraite de la description du latin indiquée sur le site du livre de Brown & Hippisley (2012)<sup>44</sup>.

**Statut des traits morphosyntaxiques** Les traits sont indiqués comme des chemins à l'intérieur de la hiérarchie morphologique.

Par ailleurs, *Network Morphology* emploie également des hiérarchies de traits indépendantes dont la structure peut être comparée à celle des hiérarchies morphologiques. Nous revenons sur ce point au chapitre 4.

44. [http://networkmorphology.as.uky.edu/sites/default/files/show\\_lat\\_mor\\_0.dtr](http://networkmorphology.as.uky.edu/sites/default/files/show_lat_mor_0.dtr) (valable au 11 juin 2013).

```

1 #show
2 ...
3
4 <mor active imperfective present indicative sg 2>
5 <mor active imperfective present indicative sg 3>
6 <mor active imperfective present subjunctive sg 1>
7 <mor active imperfective present subjunctive sg 2>
8 <mor active imperfective present subjunctive sg 3>
9 <mor active imperfective present imperative sg 2>
10 <mor active imperfective present imperative pl 2>
11 <mor active imperfective present infinitive>
12 <mor active imperfective present participle>
13
14 ...
15
16 <mor passive perfect past indicative sg 2>
17 <mor passive perfect past indicative sg 3>
18 <mor passive perfect past subjunctive sg 2>
19 <mor passive perfect past subjunctive sg 3>.

```

FIGURE 2.8 : Commande #show DATR

**La notion de défaut** En dehors du fait que l'architecture globale de *Network Morphology* repose sur la notion d'héritage par défaut, *Network Morphology* comporte une théorie complète de la notion de défaut qui différencie notamment entre les défauts exceptionnels et les défauts normaux.

Le défaut normal est le défaut inscrit dans la grammaire au niveau le plus élémentaire ; c'est l'héritage direct à partir du nœud supérieur. Ainsi, si un lexème LEX1 appartient à la catégorie des verbes et que les verbes sont spécifiés pour différencier trois temps, présent *prs*, passé *pst* et futur *fut*, alors LEX1 distinguera en vertu du défaut normal les temps présent *prs*, passé *pst* et futur *fut*.

En revanche, si dans une hiérarchie une sous-classe de verbes *C* ne différencie que le présent *prs* et le passé *pst*, alors cette distinction réduite constitue une exception par rapport au défaut normal. Il se peut toutefois qu'à l'intérieur de cette sous-classe *C*, c'est-à-dire plus bas dans la hiérarchie, se trouve un lexème LEX5 qui distingue les trois temps distingués par la catégorie verbe. Le fait de distinguer les trois temps constitue une exception par rapport au défaut à l'intérieur de la sous-classe à laquelle appartient LEX5. Mais *Network Morphology* permet de distinguer entre une exception supplémentaire s'ajoutant à la première introduite par *C* et une exception qui correspond au fait de revenir au défaut initial (distinguer les trois temps). Cette seconde exception est appelée *défaut exceptionnel* du fait que c'est une exception ajoutée sur une exception qui permet de revenir au défaut

premier, valable à la racine de la hiérarchie. Nous illustrons ces deux types de défaut par la figure 2.9.

FIGURE 2.9 : Défauts en *Network Morphology*

**Traitement des exceptions** Les exceptions sont ainsi traitées par spécification dans la hiérarchie. La figure 2.10 illustre la spécification d'une exception dans la définition de l'entrée lexicale *Xali* en archi<sup>45</sup>. Dans cet exemple les noms héritent par défaut directement du nœud supérieur NOUN. Par ailleurs, l'entrée spécifie que pour sa morphologie, *Xali* hérite du nœud N\_2. Mais pour le singulier de l'entrée *Xali*, une exception (`<mor sg> == <root sg erg> N_3_PL:<>`) spécifie que les formes doivent être construites de la manière spécifiée pour l'ergatif singulier dans le nœud N\_3\_PL. Les exceptions sont ainsi exprimées par des chemins spécifiques qui dévient du chemin par défaut dans la hiérarchie.

**Modes d'implémentation** Les implémentations de *Network Morphology* se font en DATR (Evans & Gazdar, 1989, 1996), un langage dit de réseaux d'héritage non monotone.

---

45. L'archi est une langue daghestanienne parlée par environ 1 200 personnes dans les régions montagneuses du Daghestan. Les travaux les plus complets sur cette langue sont la grammaire de Kibrik (1977) et les données récoltées au sein du projet sur l'archi par le *Surrey Morphology Group*. Les données de cet exemple sont tirées du site du livre (Brown & Hippiisley, 2012) et sont consultables en ligne à la page [http://networkmorphology.as.uky.edu/sites/default/files/ch5\\_archi.dtr](http://networkmorphology.as.uky.edu/sites/default/files/ch5_archi.dtr) (valable au 11 juin 2013).



```

1 Xali:
2   <> == NOUN
3   <mor sg> == <root sg erg> N_3_PL:<>
4   <mor> == N_2:<>
5   <gloss> == family
6   <root> == χali
7   <root sg erg> == χal.
8
9 NOUN:
10  <syn cat> == n
11  <sem cat> == non_dynamic
12  <syn> == "<mor>"
13  <syn sg abs> == "<root sg abs>".
14
15 N_2:
16  <> == N_0
17  <pl> == "<root pl>" N_2_PL:<>.
18
19 N_3_PL:
20  <> == maj MOR_CASE:<>
21  <abs> == .

```

FIGURE 2.10 : Entrée lexicale du polonais en *Network Morphology*

DATR est un langage de programmation logique qui s'appuie sur le langage Prolog. Une description en DATR est appelée *théorie*. Une théorie bien formée est une théorie qui respecte la syntaxe du langage DATR (Evans & Gazdar, 1996) et renvoie la valeur `true` « vrai » (ou *!*) lorsqu'on l'exécute. Les contraintes de bonne formation d'une description *Network Morphology* comprennent à la fois les contraintes de bonne formation du langage DATR et l'adéquation à la théorie *Network Morphology* décrite ci-dessus.

Le langage DATR repose sur une architecture Prolog dont certaines versions ne sont pas libres, mais il en existe également des versions libres. Le langage DATR qui se surimpose à Prolog est pour sa part librement disponible en ligne.

Des implémentations *Network Morphology* sont enfin librement consultables et téléchargeables sur le site du livre (Brown & Hippisley, 2012) pour dix langues différentes.

## Résumé

Le but de cette thèse est de proposer un modèle formel de la morphologie flexionnelle qui soit typologiquement fondé. Dans ce chapitre, nous avons présenté les concepts théoriques sous-tendant la conception de notre modèle formel ainsi que les deux modèles inférentiels-réalisationsnelles qui ont le plus influencé notre modèle. Dans le chapitre qui suit, nous

présentons différentes approches typologiques de la morphologie flexionnelle ainsi que les raisons qui nous ont fait choisir, parmi ces approches, le paradigme de la TYPOLOGIE CANONIQUE pour le développement de notre modèle formel.



## Qualifier la différence : la typologie canonique

The basic question of linguistic theory can be stated very simply : in what ways do natural languages differ and in what ways are they all alike ?

---

(Perlmutter, 1980, p.195)

La TYPOLOGIE LINGUISTIQUE se donne pour objectif de répondre à la question « en quoi les langues sont pareilles et en quoi elles sont différentes » (Lazard, 2003). Traditionnellement, cet objectif se traduit par deux tâches complémentaires : en étudiant un maximum de langues possibles faire l'inventaire des différences observables entre les langues et poursuivre en parallèle la « quête des universaux ou des invariants interlangues<sup>1</sup> », c'est-à-dire chercher à identifier « les constantes du langage<sup>2</sup> ».

Dans ses début, la recherche des invariants interlangues s'est en effet manifestée par la quête des UNIVERSAUX LINGUISTIQUES. Elle est notamment illustrée par l'un des textes de référence majeurs du domaine écrit par Greenberg (1963). En partant d'un échantillon de 30 langues, Greenberg fait état de 45 universaux linguistiques. Ces universaux listent essentiellement des tendances générales observables à travers ces langues, mais également quelques universaux absolus présentés comme étant valables pour toutes les langues. Dans la section des universaux morphologiques nous trouvons ainsi l'universel n° 26 indiquant

---

1. (Lazard & Moysse-Faurie, 2005, p. 11)

2. (Lazard & Moysse-Faurie, 2005, (ibid.)).

---

une tendance pour les langues affichant des « affixes discontinus » à côté d'universaux comme le n° 42 qui exprime une règle universelle conçue comme étant valable pour toutes les langues.

« 26. Dès lors qu'une langue possède des affixes discontinus, elle emploie nécessairement aussi la préfixation, la suffixation ou les deux <sup>3</sup>. »

(Greenberg, 1963, p. 69)

« 42. Toutes les langues possèdent des catégories pronominales distinguant au moins trois personnes et deux nombres <sup>4</sup>. »

(Greenberg, 1963, p. 70)

L'idée de propriétés universelles interlangues a également inspiré l'idée d'une « grammaire universelle », présente notamment dans (Chomsky, 1968) et a ainsi posé le fondement des approches générativistes. La définition des constantes du langage relève alors d'une abstraction de caractéristiques communes à toutes les langues. Il s'agit de règles communes qu'il est possible de distiller à partir des réalisations concrètes qu'emploient les langues pour les exprimer. Le succès initial du concept de grammaire universelle dans le courant principal de la linguistique formelle a eu pour conséquence son application dans d'autres disciplines voisines de la linguistique et notamment en sciences cognitives. Suite à ce succès grandissant, la notion d'universel linguistique a toutefois été vivement critiquée, comme le montrent par exemple l'article d'Evans & Levinson (2009) et les discussions qu'il a suscitées. Les auteurs y mettent l'accent sur l'illusion des universaux linguistiques tels qu'importés de (Chomsky, 1968) et sur la nécessité, notamment en sciences cognitives, de revenir aux véritables études typologiques indiquant justement la très grande diversité des langues.

Par ailleurs, l'idée de constantes du langage définissables par une abstraction des caractéristiques communes avait également été formulée en dehors du cadre de la grammaire universelle par Benveniste (1966) qui se réfère déjà à cette notion d'abstraction en employant le terme de *fonctions*.

« Ce qu'il y a de comparable dans des systèmes linguistiques complètement différents entre eux, ce sont des fonctions, ainsi que les relations entre ces fonctions, indiquées par des marques formelles. »

(Benveniste, 1966, p. 222)

Plutôt que de rechercher des universaux concrets, une approche plus satisfaisante du concept d'invariants interlangues consiste à les concevoir comme des concepts théoriques.

---

3. « 26. If a language has discontinuous affixes, it always has either prefixing or suffixing or both. »

4. « 42. All languages have pronominal categories involving at least three persons and two numbers. »

La quête de ces invariants est alors celle de la définition d'un vocabulaire descriptif suffisamment général pour rendre compte des propriétés de l'ensemble des langues étudiables et suffisamment précis pour pouvoir qualifier les phénomènes particuliers effectivement observés.

Reste que définir des constantes interlangues, mêmes abstraites, requiert l'observation de grandes quantités de données. Compte tenu des outils et des données dont dispose aujourd'hui la communauté linguistique, un inventaire de 30 langues ne saurait être suffisant<sup>5</sup>. Aujourd'hui, des inventaires conséquents sont par exemple mis à disposition sous forme de bases de données en ligne par les auteurs du *WALS*<sup>6</sup> (Dryer & Haspelmath, 2011), d'*AUTOTYP*<sup>7</sup> (Bickel & Nichols, 2002) et par le *Surrey Morphology Group*<sup>8</sup>.

La définition des constantes interlangues doit s'appuyer sur des notions suffisamment abstraites pour être cohérentes avec toutes les réalisations observées à travers les langues, comme la notion d'*accord* ou de *sujet*. Ces notions ne peuvent être des notions arbitraires, mais se doivent d'être consciencieusement forgées, et explicitement définies. Ce sont aussi ces notions qui peuvent servir de référence pour l'étude des similitudes et dissimilitudes entre les langues.

Ainsi, la comparaison des langues peut se faire selon deux méthodes complémentaires : 1) définir des tendances générales partagées en comparant directement les langues entre elles et en tentant d'extraire des points communs observables et les fondamentaux partagés par toutes ; 2) à partir de notions fondamentales conçues, de façon rigoureuse, selon la première méthode, définir un standard et comparer individuellement chaque langue à ce standard.

Tandis que la méthode consistant à définir des universaux linguistiques, adoptée notamment par Greenberg (1963), exploite immédiatement la première stratégie, d'autres approches de la typologie, et notamment la *TYPLOGIE CANONIQUE*, s'appuient essentiellement sur la seconde méthode.

Dans cette section, nous présentons brièvement les enjeux et les liens entre les différentes approches de la typologie. Dans la mesure où la suite de cette thèse s'inscrit directement dans cette approche, nous développons ensuite plus en détail la démarche adoptée par la typologie canonique, ainsi que les principales notions qu'elle manipule.

---

5. Parmi les 45 universaux stipulés par Greenberg en 1963 certains ont pu être réfutés compte tenu d'observations plus récentes de faits de langues alors non pris en compte par l'auteur. Voir par exemple Lazard & Moysse-Faurie (2005, p. 4) citant Dryer (1998, p. 301) à propos de l'universel n° 1.

6. Base de données en ligne du *World Atlas of Linguistic Structures*, accessible à l'adresse <http://wals.info/> (valable au 11 juin 2013).

7. Base de données en ligne accessible à l'adresse <http://www.spw.uzh.ch/autotyp/> (valable au 11 juin 2013).

8. *Surrey Morphology Group Database* : base de données en ligne accessible à l'adresse <http://www.smg.surrey.ac.uk/> (valable au 11 juin 2013).

## 3.1 Définir les concepts fondamentaux

Une façon d’approcher la diversité des systèmes linguistiques est de tenter d’élucider les caractéristiques qu’ils partagent. Cette tâche entamée avec la quête des universaux du langage théorisée par Greenberg (1963) passe — y compris dans les approches plus modernes nuançant le concept même d’universel linguistique — par deux approches complémentaires : 1) observer un grand nombre de données variées pour faire un inventaire des phénomènes observables ; 2) par rapport aux données disponibles, mesurer quelles sont les caractéristiques les plus partagées, et éventuellement les conditions auxquelles leur apparition est liée. En somme, l’étude de la diversité linguistique requiert simultanément une approche qualitative, pour élucider les phénomènes pertinents pour la comparaison, et une étude quantitative, pour évaluer leur importance par rapport au fonctionnement général des langues naturelles.

En un premier temps, l’étude qualitative des langues a pour but de préciser les notions traditionnelles de la grammaire afin de les doter de définitions compatibles avec la description des différentes langues observées. Un exemple de cela est la définition du passif de Siewierska (2012) que nous avons citée au chapitre 1. Le but est que chaque notion soit identifiée par des propriétés définitoires, identifiables par des tests conçus pour capter avec le plus de précision possible la généralisation permettant de relier entre elles les manifestations individuelles de cette notion à travers les langues.

### 3.1.1 Un exemple de définition conceptuelle : la reduplication

Nous présentons ici un exemple concret de définition de concepts linguistiques pour la typologie. L’exemple en question est la définition de la notion de *reduplication* par Rubino (2011)<sup>9</sup>. En guise de définition, l’auteur indique en un premier temps les propriétés générales et définitoires de la reduplication à travers les langues ; il détaille ensuite quelques exemples de reduplication observables dans les langues du monde et les fonctions possibles que peut remplir la reduplication.

Selon Rubino (2011), la reduplication se définit de manière suivante<sup>10</sup> :

« La répétition de matériau phonologique à l’intérieur d’un mot<sup>11</sup> motivée

---

9. Les données exposées dans cette section sont celles des auteurs indiqués telles que citées par Rubino (2011).

10. La traduction est la nôtre.

11. Notons que l’un des problèmes essentiels quand on définit des notions théoriques consiste en le fait que d’autres notions théoriques sont le plus souvent nécessaires pour les décrire. Une architecture théorique parfaitement solide se devrait par conséquent de définir graduellement chaque concept, en partant du plus

par la sémantique ou la grammaire est appelée **réduplication**, un mécanisme morphologique largement employé à travers les langues du monde. [...]

La **réduplication complète** consiste en la répétition d'un mot complet, d'un radical (c'est-à-dire la racine plus un ou plusieurs affixes) ou de la racine. Des exemples de réduplication complète sont

- la réduplication de mots entiers en nez-percé (langue sahaptienne ; nord-ouest des États-Unis) : té:mul « *grêle* » *vs.* temulté:mul « *neige fondue* » (Aoki, 1963, :43) ;
- la réduplication complète de la racine en tagalog (avec le préfixe marquant la catégorie verbale mag-) où le reduplicant isip est identique à la base penser : mag-isip « *penser* » *vs.* mag-isip-isip « *réfléchir sérieusement à qqch* ».

La **réduplication partielle** peut se présenter sous une grande variété de formes, depuis la simple gémiation consonantale ou l'allongement vocalique jusqu'à une copie quasi-complète de la base. En pangasinan (langue austronésienne des Philippines) par exemple, des types de réduplication divers sont employés pour former les pluriel des noms. »

L'auteur donne notamment des exemples de la manière dont la notion définie se manifeste dans diverses langues, à savoir en pangasinan, en squamish, en hunzip, en choctaw, en paumarí et en alablak.

### 3.1 Pangasinan (Rubino, 2001, :540)

CV-	toó	« <i>homme</i> »	totóo	« <i>gens</i> »
-CV-	amigo	« <i>ami</i> »	amimígo	« <i>amis</i> »
CVC-	báley	« <i>ville</i> »	balbáley	« <i>villes</i> »
Clv-	plato	« <i>assiette</i> »	papláto	« <i>assiettes</i> »
CVCV-	manók	« <i>poulet</i> »	manómanók	« <i>poulets</i> »
ce-	duég	« <i>kérabau (buffle d'eau)</i> »	deréweg	« <i>kérabaux</i> »

### 3.2 Squamish (Kuipers, 1967)

TOTALE	k°a'ǰk°a'ǰ	« <i>jouer à cache-cache</i> »	à partir de la racine	k°a'ǰ
PARTIELLE	s-λ'λ'łmut	« <i>personnes âgées</i> »	à partir du singulier	s-λ'łmu't

Réduplication affixale : préfixes, suffixes et infixes :

élémentaire pour arriver au plus complexe.



- 3.3 Hunzib (nakh-daghestanien ; Caucase oriental) : réduplication CV(C) (van den Berg, 1995, :34)  
 bat'iyab « différent » bat'bat'iyab « très différent »  
 muğáλ « après » mu.muğáλ « bien plus tard »
- 3.4 Choctaw (Muskogean ; Mississippi et Alabama) : réduplication CV interne (Kimball, 1988, :440)  
 tonoli « rouler » tononoli « rouler en avant et en arrière »  
 binili « être assis » bininili « se lever et s'asseoir »
- 3.5 Paumarí (langue arawane ; Amazonie, Brésil) : réduplication dissyllabique final (Chapman & Derbyshire, 1991)  
 a-odora-dora-bakhia-loamani-hi  
 IPL-rassembler-REDUP-souvent-vraiment-thème  
 « Nous passons notre temps à les rassembler. »

Décrire les fonctions possibles d'un phénomène comme la réduplication fait également partie de sa définition. Ainsi Rubino indique également la diversité de fonctions que peut avoir la réduplication à travers les langues. Il précise que la réduplication verbale peut dénoter la pluralité, la distribution ou la collectivité en tant qu'expression du nombre, la distribution d'un argument, le temps l'aspect (action continue ou répétée, action terminée, inchoatif), l'atténuation, l'intensité la transitivité (valence, perte du focus sur l'objet) ou encore la réciprocité. La réduplication nominale peut avoir fonction d'expression du nombre (collectif, distributif, multiplicatif, limitatif), de cas, de distribution, d'indéfinitude, de réciprocité, de taille (autant en tant que diminutif qu'en tant qu'augmentatif) ou d'associatif. Enfin, la réduplication peut également être l'expression de la morphologie dérivationnelle.

Certains de ses exemples sont résumés dans le tableau 3.1 page 152. Les segments rédupliqués sont indiqués en rouge (un mot entièrement rouge indique une réduplication totale, les autres cas sont des cas de réduplication partielle. On notera toutefois qu'il reste difficile de collecter, même pour un phénomène *a priori* aussi précis que la réduplication, des données cohérentes. En effet, si l'on observe les données répertoriées dans le tableau 3.1, il y a une continuité entre la répétition d'un segment comme dans les exemples de l'alabama, par exemple *loca/lóoca*, la répétition d'une suite de segments comme *ripok/riripok* en bukiyip et une réduplication complète comme en kayardild *kandu/kandukandu*. Néanmoins, définir cette variation à l'intérieur même de la réalisation d'une catégorie typologique fait partie du travail de mise en place du vocabulaire descriptif

essentiel à l'analyse grammaticale typologiquement fondée.

### 3.1.2 Résumé des propriétés de définitions notionnelles pour la typologie

Pour résumer, la définition de concepts utilisables en typologie doit comporter trois informations fondamentales : 1) une description générale des caractéristiques formelles d'une notion donnée ; 2) un inventaire des différentes réalisations observables ; 3) une description des fonctions du phénomène sous-jacent au concept défini.

De telles définitions permettent de regrouper en tant que réalisations d'une même catégorie typologique des phénomènes qui, malgré leur apparentes différences à travers les langues, partagent des caractéristiques formelles et/ou fonctionnelles communes. Nous pouvons alors évaluer l'ensemble des langues quant au fait de présenter un certain phénomène (cf. langues faisant usage de la reduplication *vs.* langues n'en faisant pas usage), ainsi que quant à la manière dont ce phénomène se réalise ou de la fonction qu'il remplit dans une langue donnée.

De telles comparaisons permettent ainsi des études typologiques quantitatives, comme celles qui accompagnent chaque chapitre (centré à chaque fois sur une notion particulière) dans (Dryer & Haspelmath, 2011). Pour la reduplication, Rubino (2011) indique que sur 368 langues étudiées 277 langues affichent et de la reduplication complète et de la reduplication partielle, 35 langues affichent seulement de la reduplication complète et 56 langues n'affichent aucune reduplication productive<sup>12</sup>.

## 3.2 Fondamentaux et prototypes

Nous avons vu dans la section précédente qu'il est nécessaire de procéder en un premier temps à une approche qualitative de comparaison de phénomènes observables à travers les langues du monde pour arriver à des définitions de concepts utilisables en typologie. Ces concepts ayant été définis, nous pouvons évaluer l'étendue d'un phénomène donné à travers les langues du monde. De telles études quantitatives sont notamment directement à l'origine de la définition de la notion d'universel linguistique et permettent également la définition de prototypes.

---

12. Une carte de la répartition de ces langues est accessible en ligne à l'adresse [http://wals.info/feature/27A?tg\\_format=map&v1=cd00&v2=cf6f&v3=cfff](http://wals.info/feature/27A?tg_format=map&v1=cd00&v2=cf6f&v3=cfff).

DOMAINE	FONCTION	BASE	TRADUCTION	RÉDUPLIQUÉ	TRADUCTION	LANGUE	RÉFÉRENCE
VERBE	aspect permanent	loca	« être noir (de site) »	looca	« être noir (de peau) »	alabama (muskogéen central)	(Hardy & Montler, 1988)
VERBE	pluriel	lawi	« faire un trou »	law-lawi lawa-lawi	« faire deux trous/refaire le même trou » « faire beaucoup (= plus de deux) trous »	luiséno (uro-aztèque; Cahifornie)	(Kroeber & Grace, 1960)
VERBE	intensité	su	« toucher, tenir »	susu	« toucher de partout »	bukiyip (arapesh, torricelli, Papouasie Nouvelle Guinée)	(Dobrin, 2001)
	distribution, manque de contrôle	řipok	« couper »	řřipok	« charrier »		
ADJECTIF	atténuation	kasatka	« froid »	kásatka	« frais »	alabama (muskogéen central)	(Hardy & Montler, 1988)
ADJECTIF	intensité	lamnaki	« droit »	lamnaki balak-balak xa-xabay	« assez droit » « très gros » « quelque peu inquiet »	lampung (lampungique, austronésien; Sumatra)	(Walker, 1976)
NOM	réciprocité	jokwa-t	« canards-éciers (pl.) »	balen-bales jokwa-jow	« se venger l'un l'autre » « canard-écier (SG.ABSOLU) »	ilocano	(Rubino, 2000, :84)
NOM	associatif	k'chis	« fèves »	k'č k'chis	« personne ayant de grosses fèves »	rchoukche (rchoukche-kamrchadal)	(Dunn, 1999, :132)
NOM	limitatif			sab-sábong	« quelques fleurs »	yawelmani (yokuts, pénu-tien; Californie)	(Newman, 1944)
NUMÉRAL	limitatif			wal-wal6	« seulement huit »	ilocano	(Rubino, 2011)
NUMÉRAL	distributif			ge-gel	« dix chacun, par dizaines »	santali (munda; Inde)	(Rubino, 2011)
NUMÉRAL	limitatif	asem ténel	« trois » « sept »	tal-talora asensem ténenel	« trois seulement » « trois chacun » « sept chacun »	pangasinan (austronésien; Philippines)	(Rubino, 2011)
						ao (tibéto-birman; Inde)	(Gurubasave Gowda, 1975, :39)
DÉRIVATION	adj dénominal	kandu	« sang »	kandukandu	« rouge »	kayardild (tangique; Australie)	(Evans, 1995)
DÉRIVATION	adj déverbal	lepi	« tanner, adoucir »	lepé-łpi-š	« pliable »	luiséno (uro-aztèque; Cahifornie)	(Kroeber & Grace, 1960)
DÉRIVATION	nom déverbal	giak	« envoyer »	gi-giak	« messenger »	tigak (austronésien; Papouasie Nouvelle Guinée)	(Beaumont, 1979)
DÉRIVATION	verbe déadjectival	lóm	« difficile »	lóm lom	« rendre qqch difficile »	khoekhoe (khoisan, Namibie)	(Hagman, 1977, :18)

TABLE 3.1 : Inventaire de cas de reduplication d'après Rubino (2011)

### 3.2.1 Définir des universaux linguistiques

Définir des universaux interlangues peut recouvrir deux réalités distinctes, mais néanmoins souvent complémentaires. Ainsi nous pouvons définir la notion d'universel interlangue comme un propriété partagée au même titre par toutes les langues du monde : par exemple affirmer que *toutes les langues possèdent un moyen pour exprimer la corréférence entre le locuteur et l'agent d'une action donnée* correspond à un universel absolu. Mais le terme d'universel linguistique peut également avoir un autre sens, moins restrictif, de tendances observables à travers les langues.

Parmi les 45 universaux proposés par Greenberg (1963) nous pouvons lire côte-à-côte des formulations qui relèvent de ces deux acceptions d'universel. Ainsi, l'universel 37 relève du type d'universel absolu, tandis que l'universel 1 présente une tendance générale observable à travers les langues.

« 37. Une langue n'affiche jamais moins de distinctions en genre au singulier que dans les autres nombres<sup>13</sup>. »

(Greenberg, 1963, p. 68)

« 1. Dans une phrase déclarative comportant un sujet et un objet nominaux, l'ordre dominant correspond presque toujours à celui où le sujet précède l'objet<sup>14</sup>. »

(Greenberg, 1963, p. 70)

De façon orthogonale à la distinction d'*universel absolu vs. tendance*<sup>15</sup>, (Greenberg, 1966) fait également la différence entre les *universaux implicationnels* et les *universaux non implicationnels* (Piantadosi & Gibson, 2011). Les universaux implicationnels se distinguent de ces derniers en ce qu'ils stipulent une relation (absolue ou tendancielle) entre différents phénomènes concomitants. Un exemple de ce type est l'universel 29 de Greenberg (1963), tandis que l'universel 42<sup>16</sup> illustre un cas indiscutable d'universel non implicationnel.

« 29. Toutes les langues possèdent des catégories pronominales distinguant au moins trois personnes et deux nombres<sup>17</sup>. »

(Greenberg, 1963, p. 70)

13. « 37. *A language never has more gender categories in nonsingular numbers than in the singular.* »

14. « 1. *In declarative sentences with nominal subject and object, the dominant order is almost always one in which the subject precedes the object.* »

15. Greenberg parle aussi souvent d'*universel statistique*. Nous suivons ici (Piantadosi & Gibson, 2011) en préférant le terme de tendance, moins ambigu.

16. Voir aussi la citation plus haut.

17. « 42. *All languages have pronominal categories involving at least three persons and two numbers.* »

« 29. Quand une langue affiche de la morphologie flexionnelle, elle possède toujours également de la morphologie dérivationnelle <sup>18</sup>. »

(Greenberg, 1963, p. 69)

Ainsi, nous pouvons différencier quatre types d'universaux linguistiques : (A) les universaux absolus non implicationnels, (B) les universaux absolus implicationnels, (C) les tendances non implicationnelles et (D) les tendances implicationnelles. Dans l'inventaire de Greenberg (1963), le troisième type est de loin le moins représenté (1 sur 45), suivi du premier (4 sur 45), puis du quatrième (11 sur 45) et enfin du deuxième (29 sur 45). Nous résumons cette distinction dans le tableau 3.2.

	NON IMPLICATIONNEL	IMPLICATIONNEL
ABSOLU	A : 4/45	B : 29/45
TENDANCE	C : 1/45	D : 11/45

TABLE 3.2 : Les types d'universaux chez Greenberg (1963)

Le plus grand inconvénient dans le fait de stipuler des universaux absolus est qu'ils seront invalidés dès lors que des études subséquentes permettront de montrer qu'une langue jusqu'alors non prise en compte contredit son universalité. Dans un tel cas, l'universel absolu se verra réduit à un universel tendanciel.

Une des réponses constructives à la critique de la notion d'universel linguistique formulée par Evans & Levinson (2009) a cependant été formulée par Martin Haspelmath également publiée dans (Evans & Levinson, 2009). Haspelmath y indique que bien plus que les universaux absolus, ce sont les universaux implicationnels qui méritent de retenir l'attention dans une approche moderne et éclairée de l'universalité. Le but devient ainsi d'observer des tendances de cooccurrence de propriétés plutôt que des vérités générales difficiles à vérifier. Ce dernier point est par ailleurs particulièrement souligné dans la réponse de Michael Tomasello à Evans & Levinson (2009).

Dans une approche moderne et éclairée des universaux linguistiques, nous pouvons, de la même manière que ce qui a été fait par Greenberg (1963), croiser les comptes faits dans (Dryer & Haspelmath, 2011) à l'intérieur de chaque chapitre pour en déduire des universaux implicationnels sur un inventaire de langues plus conséquent que celui étudié par Greenberg (1963). Plus nous augmentons le nombre de langues dans notre inventaire, plus nous pouvons affiner l'étude quantitative permettant de définir les universaux de type tendance et plus nous nous rapprochons des tendances réelles affichées par l'ensemble des

18. « 29. *If a language has inflection, it always has derivation.* »

langues du monde.

L'avantage des universaux de type tendance est donc qu'ils ne seront pas invalidés dès lors qu'une nouvelle langue sera décrite affichant un comportement ne s'y conformant pas. L'on peut d'ailleurs s'attendre à ce que, à partir d'un nombre significatif de langues étudiées, la description d'une nouvelle langue ne changera pas significativement les tendances générales mise à jour jusque là.

Ces universaux ainsi mis en place peuvent alors servir de repère auquel l'on peut comparer les différentes langues du monde. Une telle comparaison revient à évaluer si pour un universel donné, une langue donnée est ou non conforme à cet universel.

### 3.2.2 La notion de prototype

Dès lors que nous avons établi un inventaire de tendances générales à travers les langues, nous pouvons comparer chaque langue à chacun des universaux et évaluer combien une langue se conforme à l'inventaire complet des universaux. Une autre méthode, plus directe, serait de définir à partir des universaux établis ce qui serait, pour les langues naturelles, une *LANGUE PROTOTYPIQUE*. Cette langue prototypique aurait ainsi la caractéristique d'être conforme non seulement à l'intégralité des universaux absolus, mais de respecter également au maximum les tendances générales.

Ainsi, la notion de langue prototypique, langue qui potentiellement ne se rencontrera pas telle quelle dans les langues du monde, constitue une référence ou un standard avec lequel comparer les langues réellement étudiées et dont il serait possible de calculer le taux de déviation. Plus une langue serait différente de ce standard, plus elle serait éloignée de la notion de langue prototypique.

Notons que la notion de *PROTOTYPE* s'applique par ailleurs à des universaux tels que nous les avons décrits plus haut, mais également aux concepts fondamentaux de la description linguistique. Ainsi, Creissels (2006, p.284) utilise-t-il par exemple la notion de *VERBE D'ACTION PROTOTYPIQUE* pour définir la notion de transitivité. Dans la mesure où nous avons établi à la section 3.1.2 que la définition d'un concept de base doit s'accompagner d'un inventaire de formes de réalisation possibles de ce concept, nous pouvons également considérer la comparaison de la forme de réalisation dans une langue donnée avec la réalisation prototypique d'un concept comme un moyen de comparer entre elles les langues du monde — et les regrouper par types.

## 3.3 La morphologie naturelle

### 3.3.1 Motivations pour une morphologie naturelle

Une approche de la morphologie qui à sa manière repose sur l'idée de prototype est la MORPHOLOGIE NATURELLE développée par (Mayerthaler, 1981) et Wurzel (1984) sur le modèle de la PHONOLOGIE NATURELLE de Stampe (1973). La particularité de cette approche réside en le fait qu'elle cherche explicitement à définir la notion de FORME NATURELLE de motivations externes à la phonologie (Stampe, 1973) ou la morphologie (Mayerthaler, 1981 ; Wurzel, 1984 ; Dressler, 2000).

Comme l'indique Dressler (2000), la naturalité a essentiellement deux origines : d'une part, la notion de MARQUAGE telle que définie par Jakobson<sup>19</sup> en particulier et par l'École de Prague en général, introduisant la notion sémiotique d'ICONICITÉ de Peirce (1965) en linguistique ; de l'autre, l'adaptation directe du modèle de Stampe (1973) de PHONOLOGIE NATURELLE, introduit en morphologie par Wurzel (1984). Le concept de naturalité est un concept scalaire. Dressler indique qu'un phénomène peut avoir des réalisations plus ou moins naturelles : par exemple, le pluriel anglais *wives* de WIFE « femme, épouse » est moins naturel que le pluriel *knives* de KNIFE « couteau », mais plus naturel que le pluriel *feet* de FOOT « pied ». En simplifiant, le degré de naturalité est ici conditionné par la représentation formelle la plus simple de la valeur de pluriel, à savoir l'ajout d'un morphème du pluriel *-s*. Pour *knives* (naif *vs.* naifs) cette marque du pluriel *s* est clairement identifiable, tandis que le radical naif reste inchangé. Dans le cas de *wives* (waif *vs.* waivz), la marque est reconnaissable sous sa forme allomorphique *z*, alors que le radical subit une légère altération formelle *waiv* (voisement de la consonne finale) — tout en restant cependant parfaitement identifiable comme une variante de *waif*. Dans le cas de *feet* en revanche, la marque *s* a complètement laissé la place à un marquage SUPPLÉTIF sur le radical qui se manifeste par un changement vocalique (fut *vs.* fi:t). Pour simplifier, le marquage est ainsi d'autant plus naturel qu'il est transparent. En réalité, la notion de naturalité en morphologie est soumise à une interaction plus complexe de principes que nous résumons dans la section 3.3.2.

Les réalisations observables des concepts fondamentaux étant conçus comme soumis à un ensemble de principes, la morphologie naturelle cherche à fournir une explication plausible de ce qui conditionne formellement la structure du système morphologique d'une langue donnée. Les principes en question sont un ensemble de principes extérieurs au système morphologique, qui en tant que contraintes conditionnent la structure

---

19. Voir en particulier (Jakobson, 1965).

morphologique des langues du monde. Les différences observables à travers les langues sont conçues comme le fruit de violations de certaines de ces contraintes, différentes à chaque fois, ainsi que de leur interaction. Les principes eux-mêmes sont pour la plupart dérivés de contraintes sémiotiques, mais aussi dérivés des contraintes stipulées pour la phonologie naturelle (Stampe, 1973).

### 3.3.2 Paramètres de la morphologie naturelle

Les principes régissant la morphologie naturelle sont pour la plupart des principes dérivés directement de contraintes sémiotiques. Dressler (2000) cite ainsi les principes suivants<sup>20</sup>, issus de l'inventaire des paramètres sémiotiques proposé par Mayerthaler (1981)<sup>21</sup>.

**Iconicité** D'après les critères d'iconicité mis en place par Peirce (1965), l'échelle de naturalité comporte quatre niveaux, listés ici du plus au moins iconique pour le marquage du pluriel : ainsi, le pluriel *spoon-s* de SPOON « cuiller » où le marquage du pluriel est clairement porté par l'ajout du suffixe *-s* est appelé **DIAGRAMMATIQUE** ; le pluriel *feet* de FOOT « pied » manifestant explicitement la différence entre singulier et pluriel mais sans dissocier formellement les éléments lexicaux et grammaticaux est **MÉTAPHORIQUE**, le pluriel invariable *sheep* de SHEEP « mouton » est **NON ICONIQUE** et enfin le pluriel *hon* du francique HOND « chien » marquant le pluriel de façon soustractive est **ANTI-ICONIQUE**.

**Indexicalité** Un second paramètre de naturalité, également dérivé des principes de Peirce est celui de l'indexicalité. L'indexicalité stipule notamment que l'adjacence est préférée à la distance. Ainsi dans un cas comme celui du diminutif espagnol *pueblito* ou sa variante *pueblecito* « petit village », il est plus naturel pour l'affixe diminutif *-ito* de se coller directement à la base PUEBLO « village », cf. *puebl-ito* (lit. village-DIMINUTIF), que d'être séparé d'elle par l'interfixe *-ec-* comme dans *puebl-ec-ito* (lit. village-INTERFIXE-DIMINUTIF)<sup>22</sup>.

Une conséquence de la contrainte d'adjacence veut également qu'il y ait plus généralement une préférence pour un ordre fixe des morphèmes dans un mot donné, chaque morphème étant ainsi associé de façon parfaitement non ambiguë à une classe de position donnée à l'intérieur du mot.

20. Il est toutefois à noter que d'autres principes ou hiérarchisations de principes ont pu être proposés, cf. les remarques à ce sujet dans (Dressler, 2000).

21. Les exemples cités sont également ceux de Dressler (2000).

22. Données de Dressler (2000) citées depuis (Dressler & Merlini Barbaresi, 1994).



**Transparence** La notion de transparence recouvre à la fois la TRANSPARENCE MORPHOSÉMANTIQUE et la TRANSPARENCE MORPHOTACTIQUE.

**Transparence morphosémantique :** La transparence morphosémantique est maximale dès lors que le sens d'un mot est parfaitement compositionnel (cf. l'exemple de SPOON « cuiller » + -s (PL) = *spoon-s* (cuiller-PL) « cuillers » ci-dessus).

**Transparence morphotactique :** La transparence morphotactique est maximale dès lors que les différents constituants morphématiques s'enchaînent sans aucune altération à leur frontière. Une forme d'opacité très légère est le cas où les changements phonétiques sont directement déductibles des règles phonologiques de la langue. Un exemple de ce type est la palatalisation purement phonologique apparaissant lors de la formation du nominatif pluriel péjoratif *Polak-i* prononcé *polak'i* à partir du singulier *Polak* « polonais » (prononcé *polak*). En revanche, des altérations à la frontière entre morphèmes qui ne sont pas directement déductibles des règles phonologiques générales de la langue constituent un cas d'opacité morphotactique plus importante, cf. le nominatif pluriel normal (c'est-à-dire non péjoratif) de POLAK : dans ce cas une règle morphologique provoque une palatalisation spécifique résultant en la forme *Polac-y* (prononcé *polatsi*). Le cas le plus extrême d'opacité morphotactique est celui de la supplétion, cf. les formes du verbe TO BE « être » en anglais : *be* (INF) vs. *am* (1SG.PRS) vs. *is* (3SG.PRS) vs. *was* (SG.PST).

La transparence morphotactique conditionne par ailleurs la préférence pour l'emploi de morphes continus (plutôt que discontinus comme les infixes ou les circonfixes).

**Biunicité** La biunicité stipule l'existence d'un seul sens par forme et d'une seule forme par sens. Un exemple de naturalité selon ce paramètre est la marque du pluriel en turc. Selon Dressler, le pluriel en turc est systématiquement marqué par le suffixe *-ler* qui n'exprime lui-même aucun autre sens. L'unicité est un cas moins naturel que la biunicité ; elle est notamment représentée par le superlatif en allemand : en allemand, le superlatif des adjectifs est toujours exprimé par le suffixe *-st*, mais *-st* peut quant à lui également avoir la valeur de marque de la deuxième personne du singulier pour les verbes. Le cas le moins naturel est ainsi celui où une forme peut avoir plusieurs sens, tandis que ces sens peuvent également être exprimés par d'autres formes, cf. en latin où le suffixe *-i* peut exprimer à la fois le GEN.SG, le DAT.SG, l'ABL.SG et le NOM.PL, tandis que chacune de ces valeurs peut également être exprimée par d'autres suffixes.

À ces paramètres s'ajoute celui de la *FORME NATURELLE D'UN MOT*, dérivé de contraintes phonologiques. Dressler (2000) indique que la structure la plus naturelle pour un mot est d'être bisyllabique, ou à la limite trisyllabique. La polysyllabité (de plus de trois syllabes) comme la monosyllabité sont considérés comme des structures de mots non naturelles.

Compte tenu de l'interaction des principes décrits ci-dessus, Kilani-Schoch & Dressler (2005) décrivent donc la morphologie naturelle comme une *THÉORIE DES PRÉFÉRENCES* cherchant à expliquer « *les préférences ou asymétries récurrentes de structure dans les langues du monde [...] par des motivations extérieures au langage* » (Kilani-Schoch & Dressler, 2005, p. 17). Ces préférences affectent autant les opérations morphologiques que les constituants.

### 3.3.3 Naturalité globale et naturalité locale

Toutefois, les contraintes issues des paramètres indiqués ci-dessus ne s'appliquent pas de la même façon d'une langue à l'autre. Les langues peuvent notamment privilégier le respect de certains paramètres au détriment d'autres. Dressler (2000) précise par exemple que si le turc est hautement naturel en vertu des quatre paramètres sémiotiques, la longueur des mots contrevient à la structure naturelle de mots (bisyllabité).

La notion de naturalité spécifique à un système donné a été forgée par Wurzel (1984, 1989) pour la morphologie flexionnelle afin de capter la notion de cohérence avec un système donné quand bien même elle suppose une déviation par rapport à la naturalité universelle (au sens du respect complet de toutes les contraintes conditionnées par les paramètres définis plus haut). En guise d'illustration, Dressler (2000) donne l'exemple du latin : en vertu des contraintes de transparence et d'iconicité, on s'attendrait à ce que la façon la plus naturelle pour la formation des différentes formes casuelles d'un nom donné soit de combiner sa base<sup>23</sup> (au sens de Dressler un mot existant dans le paradigme) et le suffixe casuel approprié. Or, en latin, on constate que les formes nominales sont pour leur majorité construites soit à partir du radical assorti de la voyelle thématique appropriée (cf. *TURRIS* « *la tour* » se décline en combinant le radical du nom (ici au sens de la racine *turr* assorti de la voyelle thématique *i*) avec les différents suffixes casuels), soit directement à partir de la racine (cf. *FILIUS* « *le fils* » se décline en combinant la racine *fili* avec les différents suffixes casuels). La stratégie universellement la plus naturelle n'est donc pas la stratégie la plus employée en latin. Cela a pour conséquence que le mode de construction de forme le plus cohérent avec le système morphologique nominal du latin est la combinaison de la

23. Nous présentons ici l'exemple dans la terminologie employée par Dressler. Elle ne correspond pas nécessairement à la terminologie que nous adoptons par ailleurs dans la suite de cette thèse.

racine et des suffixes casuels. Quand bien même cette stratégie ne respecte pas les contraintes universelles de naturalité, elle est, en raison de sa cohérence avec le système de la langue, la plus naturelle pour cette langue. Nous devons donc distinguer la naturalité globale qui reflète des contraintes universelles de la naturalité locale qui reflète les contraintes de cohérence interne d'un système morphologique donné.

Enfin, le concept de naturalité est dérivé de principes extérieurs, observables notamment à travers l'acquisition plus ou moins précoce de certaines structures par de jeunes apprenants. Dressler (2000) précise par conséquent que la définition générale de la naturalité et des paramètres qui la définissent est de ce fait soumise à des changements au fur et à mesure que de nouvelles données révèlent d'autres tendances observables ou permettent de mettre à jour des variables jusqu'alors négligées.

### **3.3.4 Naturalité ou autonomie**

En somme, nous pouvons dire que la morphologie naturelle se distingue par le fait qu'elle tente de dériver la structure morphologique de structures extérieures à la morphologie. Dans une démarche typologique, cette hiérarchisation se fait par ailleurs à partir d'observations quantitatives sur les langues du monde. Par l'observation des tendances principales d'une langue donnée, elles permettent de hiérarchiser les préférences conditionnant la structure de cette langue particulière et, par l'observation des langues du monde, de hiérarchiser entre elles les préférences universelles. La naturalité se distingue ainsi du prototype simple par sa prise en compte d'informations extérieures dans la définition d'une forme naturelle.

Ainsi, les principes fondamentaux de la morphologie naturelle se nourrissent de connaissances observables à l'extérieur des systèmes morphologiques et ne suppose pour cette raison pas l'autonomie de la morphologie par rapport aux autres composantes de la langue.

L'avantage d'une telle démarche est qu'elle permet d'explorer les liens de causalité entre la langue, voire les processus cognitifs généraux, et la structure des systèmes morphologiques et offre donc une vision plus générale que celle qui serait simplement centrée sur la morphologie. Le revers de cela est néanmoins inévitablement que 1) on ne peut jamais être sûr d'avoir pris en compte l'essentiel des facteurs pertinents intervenant dans les liens de causalité explorés, 2) cette approche masque également la structuration et la cohérence internes des systèmes morphologiques. Or, dans une perspective formelle comme celle qui est envisagée dans cette thèse, ce second point est essentiel. Il nous semble notamment souhaitable de pouvoir rendre compte en un premier temps des différentes composantes

du langage (phonétique, morphologie, syntaxe, ...) avant de s'intéresser aux liens entre les composantes et aux motivations extra-linguistiques. Cette démarche est celle de la typologie canonique.

### 3.4 Fondamentaux et canons

Une approche qui en revanche repose sur une vision autonome de la morphologie est la MORPHOLOGIE CANONIQUE élaborée dans le cadre plus général de la TYPOLOGIE CANONIQUE développé au sein du *Surrey Morphology Group*, notamment par Greville G. Corbett.

Dans cette section nous présentons les caractéristiques et motivations principales de la typologie canonique. Les concepts fondamentaux de la morphologie, et en particulier de la flexion, canonique seront ensuite décrits plus en détail à la section 3.4 et les différents phénomènes non canoniques présentés à la section 3.5.

#### 3.4.1 Principes généraux de la typologie canonique

Le concept de TYPOLOGIE CANONIQUE apparaît avec les travaux de Corbett (2003) avec l'objectif de permettre une meilleure compréhension de ce qui, dans le système formel que constitue une langue, différencie un idéal hypothétique, canonique, des différentes réalisations concrètes de phénomènes observables dans les langues. Dans une telle approche, il convient de ne pas confondre SYSTÈME CANONIQUE et SYSTÈME PROTOTYPIQUE. Tandis qu'un système prototypique est une combinaison abstraite de ce qui, à travers les langues du monde, est le plus fréquent — et le plus probable —, un système canonique repose sur une idée abstraite de régularité. En ce sens, ce qui est canonique n'a pas vocation à être observable. Un système canonique, défini formellement, constitue un étalon théorique qui a pour seul objectif de servir de point de repère dont on peut différencier les phénomènes non canoniques réellement attestés (Corbett, 2007a).

Comme la notion de système prototypique ou naturel, la notion de système canonique constitue un standard auquel il est possible de confronter les systèmes des différentes langues étudiées.

En effet, comme les précédents, le concept de système canonique repose sur une définition minutieuse des concepts fondamentaux. Nous avons indiqué dans la section 3.1 que toute étude typologique repose en un premier temps sur un inventaire de notions théoriques précises, ces notions théoriques étant le fruit d'observations qualitatives des

langues permettant d'identifier l'inventaire des notions théoriques nécessaires.

Dans une démarche canonique, il convient ensuite pour chaque concept (par exemple le concept de sujet, de flexion, d'accord, de cas, de nombre, ...), de définir ce que serait sa réalisation canonique. Une RÉALISATION CANONIQUE se définit par une RÉGULARITÉ MAXIMALE pour exprimer une AMBIGÜITÉ MINIMALE. Ainsi un système canonique correspond à un système qui pour chaque notion intervenant dans la constitution d'un système linguistique la réalise avec une parfaite régularité et une ambiguïté nulle. Il est évident que cette combinaison a peu de chance d'être réalisée simultanément pour toutes les notions dans une langue réellement observable. Néanmoins, cette régularité parfaite constitue un standard immuable auquel il est possible de confronter chaque système linguistique, les différences dans les déviations observables par rapport au canon permettant ainsi de comparer les langues entre elles de façon indirecte mais objective.

La différence avec les standards que constituent le système prototypique d'une part et le système maximalement naturel d'autre part réside essentiellement dans le fait que ces deux systèmes font intervenir des informations extérieures à la seule définition conceptuelle.

La notion de prototype sera construite en identifiant par des observations quantitatives quelles sont les réalisations les plus communes des notions définies au préalable. Ces réalisations peuvent alors être considérées comme les réalisations prototypiques de ces notions et permettent ainsi de faire l'inventaire des universaux observables. De ces universaux il est alors possible de déduire un système prototypique théorique qui peut ensuite servir de standard pour comparer les réalisations effectives observées. Plus proche de la réalité des langues, ce type de standard présente néanmoins le défaut de ne pas être fixe, mais de dépendre des données à disposition au moment où l'on effectue les observations quantitatives et d'être pour cette même raison susceptible de varier dès lors que de nouvelles données sont prises en compte.

La notion de système naturel, quant à lui, se nourrit d'informations extralinguistiques et présente à ce titre les défauts que nous avons indiqués en 3.3.4. Il s'agit par ailleurs d'une approche fonctionnelle (Kilani-Schoch & Dressler, 2005, p. 27) cherchant à établir dans quelle mesure une forme donnée correspond la fonction qu'elle a effectivement dans la phrase. Les deux fonctions principales étant LA FONCTION COMMUNICATIVE (Dressler, 1990) et la FONCTION COGNITIVE (Mercer, 2000 ; Carruthers, 2002). En dehors des intérêts que représente une telle recherche de motivation extérieure des structures observables en morphologie, il nous semble que le développement d'un modèle formel gagnera en solidité s'il se concentre en un premier temps à rendre compte entièrement de la structure interne d'une composante de la langue. Une fois cette tâche, en soi déjà

suffisamment vaste, accomplie pour chacune des composantes du langage (phonologie, morphologie, syntaxe, ...), il sera également plus facile d'explorer les liens qu'entretiennent entre elles ces composantes, ainsi que les liens existant entre les composantes du langage et des motivations extérieures (évolution diachronique du système, interaction avec les systèmes voisins, motivations cognitives, communicatives et contextuelles, ...).

Le schéma de la figure 3.1 récapitule la conception des trois types de standards que constituent un système prototypique, un système maximalement naturel et le système canonique. Les flèches bleues dans le schéma représentent le processus de conception propre à un système prototypique, les flèches vertes celui d'un système maximalement naturel et les flèches rouges la conception du système canonique<sup>24</sup>.

FIGURE 3.1 : Conception de standards en typologie

Parmi les trois types de standards, seul le système canonique constitue un repère immuable. Les prototypes peuvent varier en fonction de nouvelles données prises en compte. Un système naturel est soumis à la prise en compte de données extérieures, de sorte que l'identification de nouvelles variables de la naturalité peut changer l'équilibre entre les contraintes, voire en faire intervenir de nouvelles, ayant pour résultat la conception d'un nouveau standard de système naturel.

Dans la mesure où le but du travail développé dans cette thèse est de mettre en place un modèle formel pour rendre compte de la flexion en morphologie, le système dont nous avons besoin est un système soumis à une variation minimale. Nous avons donc opté pour la démarche canonique, plus immuable et de ce fait plus adapté pour servir de fondement

---

24. Parmi les trois processus, le plus complexe est celui définissant un système maximalement naturel.

à l'architecture formelle mise en place et décrite au chapitre 4.

### 3.4.2 Ontologies et canons

Le but de la typologie est par ailleurs également de faire l'inventaire des structures et phénomènes existants pour en extraire des intuitions sur quelles sont les structures linguistiques possibles. Cette réflexion sur ce qui constitue ou non une structure possible pour une langue naturelle est résumée comme suit par Hjelmslev (1966)<sup>25</sup> :

« Une typologie linguistique exhaustive est la tâche la plus grande et la plus importante qui s'offre à la linguistique. [...] En fin de compte, sa tâche est de répondre à la question : quelles structures linguistiques sont possibles, et pourquoi telles structures sont-elles possibles quand d'autres ne le sont pas ? Ce faisant, elle doit, plus qu'aucune autre espèce de linguistique, s'approcher de ce qu'on pourrait appeler le problème de la nature du langage. »

(Hjelmslev, 1966, p. 128-129)

La problématique des structures possibles a également été explorée au sein du courant de la typologie canonique. En typologie canonique, cette question repose essentiellement sur la notion d'ONTOLOGIE<sup>26</sup>.

L'article de Corbett (2007a) « *Canonical Typology, suppletion and possible words* » traite le problème des structures possibles et impossibles de la façon suivante : En spécifiant pour chaque notion sa réalisation maximale régulière pour une ambiguïté *zéro*, Corbett indique qu'il est également possible de prévoir toutes les dimensions par lesquelles les réalisations attestées et observables sont susceptibles de s'éloigner du canon. En d'autres termes, les dimensions de la régularité et de l'irrégularité une fois établies, il est possible de prévoir, de façon théorique, toutes les possibilités de réalisations potentiellement observables dans des données réelles.

La démarche canonique, par son approche purement théorique, mais reposant de façon explicite sur une définition minutieuse des concepts, a ainsi pour objectif de définir un espace théorique des possibilités à l'intérieur duquel le linguiste typologue devrait s'attendre à observer des réalisations concrètes.

Une définition théorique de ce qui, compte tenu des concepts fondamentaux définis, devrait être potentiellement observable dans les langues du monde est également un

---

25. Cité par Lazard (2003).

26. Le terme « ontologie » est ici à prendre dans son acception philosophique première : « Partie de la philosophie qui a pour objet l'étude des propriétés les plus générales de l'être, telles que l'existence, la possibilité, la durée, le devenir. », définition du TLFi, <http://www.cnrtl.fr/lexicographie/ontologie> (valable au 11 juin 2013).

point de départ extrêmement utile pour la conception d'un modèle formel. En effet, pour notre travail visant à établir un modèle formel de la morphologie flexionnelle utilisable pour la description et la modélisation d'un maximum de langues, cette approche présente l'avantage de prévoir de façon théorique l'intégralité des phénomènes susceptibles d'apparaître et que, par conséquent, notre modèle formel se doit de pouvoir capter.

À l'inverse, une démarche reposant sur une approche ontologique permet également d'exclure ce qui *a priori* n'est pas susceptible d'être réalisé par les langues du monde. Un modèle conçu à partir d'une telle théorie sera contraint à ne permettre la modélisation que des seuls phénomènes ontologiquement valides. Cela offre un double avantage pour l'analyse des structures morphologiques des langues : 1) La description des langues par les seules catégories définies comme ontologiquement valides permet de plus facilement cerner les particularités d'une langue, l'ensemble des structures possibles étant d'emblée restreint. 2) Quand bien même le rôle premier d'un modèle formel est, de notre point de vue, de fournir une description cohérente des structures modélisées, la formalisation d'une langue dans un modèle ontologiquement contraint permet<sup>27</sup> d'amorcer le processus d'explication de ces structures. En effet, si un modèle formel repose sur un modèle ontologique bien conçu, ne pas pouvoir formaliser une structure attestée dans ce modèle devrait pouvoir indiquer que la tentative de formalisation n'est pas la bonne. La bonne description des phénomènes observés ne peut pas être celle que l'on pensait, indiquant que soit une erreur s'est glissée dans la modélisation, soit des données supplémentaires essentielles n'ont pas été prises en compte, alors que ce sont elles qui permettraient de modéliser le système complet de façon ontologiquement valide. Un modèle ainsi contraint peut par conséquent servir de garde-fou à la description, tandis qu'un modèle moins contraint perdrait cet avantage.

### 3.4.3 La morphologie canonique

Notre travail s'intéresse à l'application de la typologie canonique à la morphologie, autrement dit à la MORPHOLOGIE CANONIQUE. En particulier, notre but est de présenter un modèle formel de la flexion qui s'appuie sur la notion de FLEXION CANONIQUE. Dans cette section nous donnons donc un rapide aperçu des caractéristiques principales de la morphologie canonique. Les phénomènes canoniques et non canoniques de la flexion

---

27. Il va de soi que le modèle ontologique est hautement dépendant de la pertinence et de la minutie avec lesquelles les concepts fondamentaux ont été définis. Il est par conséquent essentiel 1) de mettre en place des définitions précises, notamment non circulaires, des concepts sous-jacents ; 2) de veiller en continu à ce qu'elles restent valables quand bien même de nouvelles données pourraient être découvertes. Dans le cas contraire, une redéfinition des concepts fondamentaux concernés s'impose qui peut à son tour influencer les autres définitions et, par conséquent, le modèle ontologique.



seront décrits plus en détail en section 3.5.

Notons en un premier temps que l'approche canonique de la morphologie se fait dans un cadre lexicaliste lexématique. Elle adopte également une vision autonome de la morphologie, notamment au sens de Aronoff (1994) et Brown & Hippisley (2012)<sup>28</sup>.

L'approche standard de la morphologie canonique distingue par ailleurs les notions de RADICAL et d'EXPOSANT discutées au chapitre 2. Selon ces définitions un mot est constitué d'un radical et d'exposants. Un exposant est la partie du mot qui exprime les traits morphosyntaxiques réalisés par le mot ; le radical correspond au matériau proprement lexical qui subsiste une fois que l'on a retranché les exposants d'un mot donné.

L'approche canonique distingue en général clairement entre opérations morphologiques flexionnelles et constructionnelles. Le tableau 3.3 page 168 résume les points principaux indiqués par Brown & Hippisley (2012) par lesquels la flexion canonique se distingue de la dérivation canonique. Tandis que la flexion canonique concerne la façon canonique de construire le paradigme d'un lexème donné (c'est-à-dire l'ensemble de ses formes possibles), le concept de dérivation canonique recense l'ensemble des propriétés intervenant canoniquement dans la création de nouveaux lexèmes par dérivation à partir d'un lexème existant dans le lexique. D'après Brown & Hippisley, les différences entre les deux opérations morphologiques concernent les sept points suivants :

1. Le but de la flexion canonique est de construire les formes d'un lexème donné, le but de la dérivation canonique est de créer un nouveau lexème à partir d'un lexème par l'emploi de règles de création lexicales dérivationnelles ;
2. La flexion répond à un besoin de la syntaxe et est en cela déterminée par elle. Elle permet notamment de réaliser les formes que requièrent les règles d'accord (cf. en français, former le pluriel d'un verbe à employer dans une phrase où le sujet est au pluriel : « *Les koalas (PL) dorment (PL) 75% de leur temps* »). La dérivation est un processus lexical indépendant de la syntaxe.
3. Dans la mesure où la morphologie flexionnelle produit des formes requises par la syntaxe, chaque forme, en tant qu'elle réalise l'une de ces requêtes, est obligatoirement définie comme le résultat d'une fonction syntaxique. Il suit de (3) que cela ne vaut pas pour la dérivation.
4. La morphologie flexionnelle est entièrement productive : toutes les formes productibles à partir des règles flexionnelles existantes sont canoniquement effectivement produites par ces règles. En dérivation, certaines règles existantes ne sont pas appli-

---

28. Un argumentaire en faveur d'une morphologie autonome a été donné au le chapitre 2 de cette thèse.

cables à certains lexèmes et ne produisent donc aucun résultat : les règles dérivationnelles ne sont pas entièrement productives. Ainsi, quand bien même il existe une règle dérivationnelle permettant de construire un adjectif en *-ique* à partir d'un nom, cette règle n'est pas applicable à toutes les bases nominales : par exemple, il est possible de construire *métallique* à partir de *métal*, mais il n'existe pas de formation en *-ique* à partir de *plâtre*<sup>29</sup> quand bien même cette forme serait *a priori* possible.

5. L'application d'une règle flexionnelle donne toujours un résultat sémantiquement prédictible<sup>30</sup>. La part lexicale apportée par le lexème et la part grammaticale (réalisation de traits morphosyntaxiques particuliers, cf. *pluriel, présent etc.*) se combinent de façon entièrement transparente. En revanche, l'application d'une règle dérivationnelle peut donner lieu à des glissements sémantiques, parfois importants. Ainsi, par exemple, la combinaison de la règle dérivationnelle permettant de construire un nom d'agent à partir d'un verbe à l'aide du suffixe *-eur*, telle qu'illustrée par le couple *courir vs. coureur*, peut s'accompagner de glissements sémantiques, cf. *dormir vs. dormeur (crustacé)*.
6. Les traits syntactico-sémantiques d'un lexème (catégorie syntaxique, sens du lexème, ...) ne varient pas en flexion, mais ils sont changés par la dérivation (cf. *dormeur* n'a plus le sens d'un individu sommeillant, ni la catégorie syntaxique « verbe » de sa base *dormir*).
7. Enfin, canoniquement, les marques des règles flexionnelles sont ajoutées après les marques dérivationnelles. En effet, dans la mesure où la flexion sert à produire les formes de lexèmes existants, y compris de lexèmes eux-mêmes dérivés, elle prend en entrée un lexème existant, quand bien même il est le résultat d'une règle de création lexicale (dérivationnelle). L'application de la flexion constitue l'achèvement de la construction d'une forme<sup>31</sup>.

#### 3.4.4 La flexion canonique

Le concept de FLEXION CANONIQUE définit ce qui, dans la constitution d'un paradigme, différencie l'idéal hypothétique, canonique et parfaitement régulier des différents phénomènes flexionnels irréguliers observables dans les langues. Dans son article

29. La seule occurrence de *plâtrique* sur Google est (en dehors des fautes d'orthographe pour *pratique* ou *plastique*) une occurrence faisant référence à l'objet *plâtre* et non pas la matière (valable au 11 juin 2013).

30. Voir aussi (Wurzel, 1989, p. 36).

31. On note que de ce point de vue certaines formes fléchies comme les participes ne sont pas canoniques. En latin, par exemple, les participes peuvent à leur tour être fléchis en nombre et en cas.

PARAMÈTRE	FLEXION	DÉRIVATION
1 But	<i>construire la forme spécifique d'un lexème qu'exige un contexte syntaxique particulier</i>	<i>construire un nouveau lexème à partir d'un lexème existant</i>
2 Déterminisme syntaxique	<i>déterminée par la syntaxe</i>	<i>non déterminée par la syntaxe</i>
3 Caractère obligatoire	<i>fonction obligatoire</i>	<i>fonction non obligatoire</i>
4 Productivité	<i>entièrement productif</i>	<i>non entièrement productif</i>
5 Transparence	<i>transparent</i>	<i>pas toujours transparent</i>
6 Héritage des traits fondamentaux	<i>tous les traits fondamentaux sont hérités</i>	<i>seul un nombre limité de traits fondamentaux est hérité</i>
7 Ordre d'exposition	<i>après l'exposant dérivationnel ; termine un mot</i>	<i>avant l'exposant flexionnel ; ne termine pas nécessairement un mot</i>

TABLE 3.3 : Différences entre flexion et dérivation canoniques selon Brown & Hippiusley (2012, p. 37)

de 2007a, Corbett définit une liste de critères permettant d'identifier un paradigme canonique. Ces critères sont résumés dans le tableau 3.4. En particulier, la flexion canonique se distingue par les points suivants :

1. Chaque forme du paradigme présente la même structure. Ainsi, si les formes sont construites par la combinaison d'un radical avec un ou plusieurs suffixes (cf. les verbes latins présentés au chapitre 1 : *am-o*, *ama-s*, ...), toutes les formes seront construites ainsi. Cela vaut autant pour la comparaison des formes d'un même lexème au sein de son paradigme que pour la comparaison avec les formes des paradigmes de lexèmes différents.
- 2a. Les radicaux des formes d'un même lexème sont toujours les mêmes. En ce sens, le verbe CHANTER est canonique en ce qu'il présente un radical unique *chant-* fāt pour toutes ses formes.
- 2b. Canoniquement, les radicaux de deux lexèmes distincts sont nécessairement différents. L'homonymie n'est pas canonique.
- 3a. Le matériau flexionnel (les exposants) est différent d'une case à l'autre au sein d'un même paradigme ;

3b. mais il est identique d'une case d'un paradigme à celle d'un autre paradigme réalisant les mêmes traits.

Le critère (3) rejoint le critère de bi-unicité en morphologie naturelle : chaque exposant exprime une et une seule valeur et chaque valeur est exprimée par un et un seul exposant.

4. Le résultat de (2) et de (3) est que toutes les formes d'un même paradigme sont différentes de par leur matériau flexionnel, tandis que les contraintes sur les radicaux exprimées en (2) garantissent qu'il n'y a pas deux formes de paradigmes différents qui pourraient être identiques.

Grâce aux contraintes (1), (2a) et (3b) les paradigmes canoniques font ainsi preuve d'une régularité parfaite. Les contraintes de bi-unicité exprimées par (2b) et (3a) garantissent de surcroît que cette régularité parfaite s'accompagne d'une absence complète d'ambiguïté, telle qu'exprimé par le critère (4) : toutes les formes de la langue sont différentes, il n'existe aucune homonymie.

En réalité, il serait ici également nécessaire de poser une contrainte sur le découpage des formes en radical et exposant. Tel quel, l'homonymie reste une possibilité pour les formes complètes appartenant aux paradigmes de lexèmes de catégories morphosyntaxiques différentes (les lexèmes d'une même catégorie partageant canoniquement les mêmes exposants (cf. le critère n° 3). Une séquence phonique *ababa* pourra être identique à une autre quand bien même la première se découperait en radical *aba* + exposant *ba*, tandis que la seconde serait formée de *a* + *baba*. Un exemple d'homonymie de ce type en français est illustré par le couple *av-i-ons* av-j-õ (VERBE : 1PL.PST.IPFV.IND) et *avion-s* avjõ (NOM : PL).

	COMPARAISON ENTRE CASES D'UN PARADIGME	COMPARAISON ENTRE LEXÈMES
COMPOSITION/STRUCTURE	<i>identiques</i>	<i>identiques</i>
MATÉRIAU LEXICAL ( $\approx$ radicaux)	<i>identiques</i>	<i>différents</i>
MATÉRIAU FLEXIONNEL ( $\approx$ flexions)	<i>différents</i>	<i>identiques</i>
RÉSULTAT ( $\approx$ formes fléchies)	<i>différents</i>	<i>différents</i>

TABLE 3.4 : Critères pour la flexion canonique selon Corbett (2007a).

Le critère (1) imposant à toutes les cases de paradigmes canoniques d'afficher la même structure contient par ailleurs implicitement deux contraintes supplémentaires que nous

ajoutons donc explicitement aux critères de Corbett. Ces critères supplémentaires sont indiqués dans le tableau 3.5 complétant le tableau 3.4. Ils assurent le fait que toutes les cases possibles d'un paradigme donné soient remplies (critère de COMPLÉTUDE, cf. le critère n° 4 de productivité dans les critères pour la flexion canonique chez Brown & Hoppisley (2012)), et que ce remplissage des cases se fasse à chaque fois par une forme unique par case (critère d'UNICITÉ). Dans un paradigme canonique, il n'y a ni cases manquantes, ni doublons dans une même case <sup>32</sup>.

FLEXION CANONIQUE	
COMPLÉTUDE	<i>Un paradigme canonique ne comporte aucune case vide.</i>
UNICITÉ	<i>Chaque lexème dispose d'au plus une forme pour exprimer chacun des paquets de traits morphosyntaxiques pertinents pour sa catégorie (voir (Thornton, 2011)).</i>

TABLE 3.5 : Critères supplémentaires pour la flexion canonique

Enfin, les critères de Brown & Hoppisley (2012) indiquent que la fonction première de la morphologie flexionnelle est de produire les formes requises par la syntaxe : elle lui fournit des formes exprimant les traits morphosyntaxiques requis pour une construction syntaxique donnée. De ce point de vue, chaque case d'un paradigme peut être conçue comme l'expression d'un paquet de traits morphosyntaxiques donné distinct de chacun de ceux réalisés par les autres cases. Chaque couple (paquet de traits morphosyntaxiques, forme) est ainsi unique, garantissant une ambiguïté nulle pour toute forme susceptible d'être employée en syntaxe. Nous pouvons donc ajouter un paramètre supplémentaire à la définition des paradigme canoniques ci-dessus :

1. Chaque case est l'expression unique d'un paquet de traits morphosyntaxiques donné.
2. Chaque paradigme appartenant à la même classe morphosyntaxique de lexèmes affiche le même inventaire de cases. Il exprime donc par des exposants formellement identiques (cf. critère n° 3 de Corbett (2007a)) les mêmes traits que les autres paradigmes de la même classe.
3. Il suit des deux critères ci-dessus que tous les lexèmes d'une même catégorie morphosyntaxique possèdent également le même nombre de cases dans leur paradigme.

32. Voir les phénomènes de défektivité et surabondance décrits en section 3.4 pour des cas de non-respect de ce critère.

Nous avons établi un inventaire de critères définissant la flexion canonique. Toute déviation par rapport à ces critères correspond à des phénomènes non canoniques particuliers que nous présentons, illustrés par des exemples de langues variées, à la section 3.5.

### 3.5 Phénomènes flexionnels canoniques et non canoniques

Un système canonique au sens de Corbett (2007a) est un concept abstrait. Les langues du monde affichent chacune à leur manière des propriétés qui s'éloignent du canon. Ainsi, la violation plus ou moins importante de chacun des critères de canonicité constitue la signature d'irrégularité faisant la particularité d'une langue donnée. Le respect de l'intégralité de ces huit critères définit un système canonique. Les différentes violations de la canonicité sont autant de réalisations de phénomènes non canoniques.

Nous rappelons ici les critères définissant les paradigmes canoniques. La figure 3.2 donne une illustration de ce à quoi correspondraient canoniquement les paradigmes de deux lexèmes distincts. Il s'agit naturellement de données artificiellement construites, mais qui nous serviront, tout au long de cette section, à illustrer les phénomènes non canoniques par la manière dont ils affecteraient ces paradigmes. À chaque fois, nous illustrerons également les phénomènes par des exemples réels.

1. Toutes les formes sont construites par la même structure : cf. dans l'exemple de la figure 3.2, RADICAL (LO ou FA) + exposant *suffixe*. (STRUCTURE RÉGULIÈRE)
2. Les radicaux d'un même lexème sont toujours identiques, mais distincts de ceux de l'autre lexème. (BIUNICITÉ DU RADICAL)
3. Si une case du paradigme du LEXÈME 1 exprime les mêmes traits morphosyntaxiques que la case correspondante dans le paradigme du LEXÈME 2, elles comportent le même exposant, à savoir dans la figure 3.2, le même suffixe. (BIUNICITÉ DES EXPOSANTS)
4. Ainsi, toutes les formes des deux paradigmes sont différentes. (NON-AMBIGUÏTÉ DES FORMES)
5. Toutes les cases du paradigme sont remplies. Il n'y a pas de case vide. (COMPLÉTUDE)
6. Chaque case comporte au plus une forme. (UNICITÉ)
7. Chaque case est l'expression d'un paquet de traits morphosyntaxiques spécifique. Il n'y a pas deux cases exprimant le même paquet de traits morphosyntaxiques, cf.

la colonne traits dans l'exemple de la figure 3.2. (NON-REDONDANCE PAQUETS DE TRAIT MORPHOSYNTAXIQUES)

8. Chaque paradigme appartenant à des lexèmes de la même catégorie morphosyntaxique<sup>33</sup> présente le même inventaire de cases, c'est-à-dire un ensemble de cases contenant chacune une forme unique réalisant un des paquets de traits morphosyntaxiques exprimés par cette catégorie. (INVENTAIRE STABLE DES PAQUETS DE TRAIT MORPHOSYNTAXIQUES)

Traits	LEXÈME 1		LEXÈME 2	
1SG	LO	-ma	FA	-ma
2SG	LO	-sa	FA	-sa
3SG	LO	-ta	FA	-ta
1PL	LO	-mo	FA	-mo
2PL	LO	-so	FA	-so
3PL	LO	-to	FA	-to

FIGURE 3.2 : Paradigmes canoniques

Ayant défini les critères permettant de définir la FLEXION CANONIQUE, nous pouvons désormais décrire et illustrer neuf PHÉNOMÈNES NON CANONIQUES, en indiquant notamment pour chacun les critères de canonicité qu'ils violent.

Les phénomènes non canoniques que nous présentons sont regroupés sous trois catégories principales : les phénomènes concernant la RÉALISATION CANONIQUE DES FORMES (critères 1 à 4), les phénomènes concernant le REMPLISSAGE CANONIQUE DES CASES du paradigme (critères 5 et 6) et les phénomènes concernant l'EXPRESSION CANONIQUE DES TRAIT MORPHOSYNTAXIQUES (critères 7 et 8).

### 3.5.1 Non-canonicité de la forme

#### 3.5.1.1 Supplétion

Parmi les critères définitoires de la flexion canonique, Corbett (2007a) précise qu'à l'intérieur d'un paradigme d'un lexème donné, le radical reste inchangé, c'est-à-dire identique d'une forme à l'autre (critère n° 2). Il précise également que deux formes de deux paradigmes différents réalisant les mêmes traits morphosyntaxiques affichent canoniquement le même exposant (critère n° 3). La SUPPLÉTION est un phénomène défini par le non-respect du critère n° 2, et éventuellement du critère n° 3.

33. Par exemple nom, verbe, adjectif,...

Parmi les phénomènes qualifiés de supplétion, on différencie deux types : la SUPPLÉTION DE RADICAL (violation du seul critère n° 2) et la SUPPLÉTION DE FORME (Boyé, 2006). Les deux types de supplétion se caractérisent par la présence d'une forme inattendue dans le paradigme ; dans le cas de la supplétion de radical, elle ne remplace que la partie radicale d'une forme, dans le cas de la supplétion de forme, elle affecte la forme dans son ensemble.

**3.5.1.1.1 Supplétion de radical** La SUPPLÉTION DE RADICAL peut recouvrir plusieurs définitions. La plus simple est de dire qu'il s'agit d'un cas où, à l'intérieur d'un paradigme donné, les radicaux varient, alors que, d'après le critère n° 2 définissant la flexion canonique, ils devraient être stables. Notons que l'alternance des exposants reste celle attendue dans le paradigme en question selon le critère n° 3. La figure 3.3 illustre un cas de supplétion de radical pour le paradigme du lexème 1. Les formes marquées en rouge correspondent aux formes de radical variant du LO attendu.

Une autre définition, plus restreinte, de la supplétion consiste à dire que la supplétion est une forme extrême d'ALLOMORPHIE DE RADICAL. Selon cette définition, toute variation du radical est un cas d'allomorphie. Seules les variations n'ayant plus rien en commun avec les autres formes sont des cas de supplétion à proprement parler<sup>34</sup>.

Traits	LEXÈME 1		LEXÈME 2	
1SG	LO	<i>-ma</i>	FA	<i>-ma</i>
2SG	LO	<i>-sa</i>	FA	<i>-sa</i>
3SG	RI	<i>-ta</i>	FA	<i>-ta</i>
1PL	LO	<i>-mo</i>	FA	<i>-mo</i>
2PL	MU	<i>-so</i>	FA	<i>-so</i>
3PL	GA	<i>-to</i>	FA	<i>-to</i>

FIGURE 3.3 : Supplétion de radical

Un exemple évident de supplétion de radical en français est fourni par le verbe *aller* pour lequel les descriptions indiquent jusqu'à quatre radicaux différents ; *all-*, *v-*, *i-* et *aill-*.

**3.5.1.1.2 Supplétion de forme** La SUPPLÉTION DE FORME concerne des cas où une forme complète d'un paradigme est remplacée par une forme inattendue. En plus de la contrainte n° 2, la supplétion de forme viole donc également la contrainte n° 3, puisque l'exposant attendu pour la réalisation d'un certain paquet de traits morphosyntaxiques n'est plus identifiable, cf. les exemples en rouge dans la figure 3.4.

34. Dans le modèle mis en place au chapitre 4 cette différence sera marquée par le partage de l'information entre lexique et grammair.



Traits	LEXÈME 1	LEXÈME 2
1SG	LO <i>-ma</i>	FA <i>-ma</i>
2SG	<i>piu</i>	FA <i>-sa</i>
3SG	<i>twit</i>	FA <i>-ta</i>
1PL	LO <i>-mo</i>	FA <i>-mo</i>
2PL	<i>kwi</i>	FA <i>-so</i>
3PL	<i>pip</i>	FA <i>-to</i>

FIGURE 3.4 : Supplétion de forme

Un exemple de supplétion de forme en français est donné dans Bonami & Boyé (2002) pour le verbe *être* au présent de l'indicatif (cf. tableau 3.6) : la forme *sommes* de la première personne du pluriel, n'affiche par exemple plus l'exposant attendu que serait le *-ons*.

	SINGULIER	PLURIEL
P1	<i>suis</i>	<i>sommes</i>
P2	<i>es</i>	<i>êtes</i>
P3	<i>est</i>	<i>sont</i>

 TABLE 3.6 : Supplétion de formes dans le paradigme du verbe *être*

**3.5.1.1.3 Distribution des formes supplétives** La DISTRIBUTION DES FORMES SUPPLÉTIVES peut être plus ou moins transparente, c'est-à-dire plus ou moins fortement refléter des variations dans les traits morphosyntaxiques exprimés par des formes données. Ainsi, les langues iraniennes font usage d'alternance de radicaux pour exprimer une opposition entre les temps du passé et les temps du présent. Le persan, illustré au tableau 3.7, utilise le radical RAD1 pour les temps du présent et le radical RAD2 pour les temps du passé. Cette distribution peut ainsi paraître parfaitement transparente. Notons cependant que ces radicaux apparaissent également dans d'autres formes, pour lesquelles leur distribution est nettement moins transparente. Ainsi, le RAD2 est également employé pour former les infinitifs, tandis que le RAD1 est aussi employé dans les formes impératives.

LEXÈME	TRAD.	RAD1	RAD2
ÂRÂSTAN	« orner »	ârâst	ârâ
ÂMUXTAN	« apprendre »	âmuxt	âmuz
RAQSIDAN	« danser »	raqsid	raqs

TABLE 3.7 : Radicaux du présent (RAD1) et du passé (RAD2) en persan

FORMATION DU RADICAL DU PASSIF	LEXÈME	RAD1	RADICAL DU PRÉSENT PASSIF	TRADUCTION
RAD1	KUŠTIN	kuš	kuš-rê	« tuer »
RAD2	ÛTIN	l'ë	ût-rê	« dire »
RAD2	BISTIN	bîe	bîs-rê	« entendre »
RAD1 MOINS VOY. FIN.	KIRDIN	ke	k-rê	« faire »
RAD1 MOINS VOY. FIN.	DAN	de	d-rê	« donner »
AUTRE	XWARDIN	xo	xû-rê	« manger »
AUTRE	GIRTIN	gir	gîr-rê	« prendre »

TABLE 3.8 : Radicaux irréguliers du passif en kurde sorani

Le kurde sorani<sup>35</sup> s'éloigne un peu plus de cette transparence. Comme le montrent les grammaires de référence McCarus (1958) ; Thackston (2006) les formes du passif sont généralement dérivées du RAD1. Cependant, comme le montre le tableau 3.8, il existe des verbes pour lesquels le RAD2 sert de base à la formation des passifs, tandis que d'autres verbes encore ont des radicaux du passif complètement irréguliers, à savoir le RAD1 privé de sa voyelle finale ou encore d'autres formes totalement irrégulières (Walther, 2011a ; Walther & Sagot, 2011).

La supplétion de radicaux peut également se montrer plus généralisée. Ainsi, dans Bonami & Boyé (2003), les auteurs montrent pour la flexion verbale du français qu'il existe jusqu'à douze combinaisons de traits morphosyntaxiques possibles qui se réalisent au moyen de radicaux spécifiques. Ces combinaisons de traits correspondent à une partition du paradigme verbal induite par ce que les auteurs désignent par le terme ESPACE THÉMATIQUE ; les cases d'un espace thématique sont reliés entre elles par des règles de DÉPENDANCE THÉMATIQUE.

### 3.5.1.2 La notion de classe flexionnelle

Selon le critère n° 2 définissant la flexion canonique, deux cases appartenant à deux paradigmes différents comportent nécessairement le même exposant si elles expriment le même paquet de traits morphosyntaxiques. La notion de CLASSE FLEXIONNELLE contrevient à cette contrainte en ce qu'elle définit des classes de lexèmes qui se distinguent par le fait qu'elles n'emploient pas le même inventaire d'exposants pour exprimer les paquets de traits morphosyntaxiques en question.

35. Langue iranienne occidentale parlée essentiellement à la frontière irano-irakienne.

Ainsi, comme le montrent les exposants dans l'exemple de la figure 3.5, les lexèmes 1 et 2 appartiennent à une même classe flexionnelle, distincte de celle des lexèmes 3 et 4.

Traits	LEXÈME 1		LEXÈME 2	
1SG	LO	-ma	FA	-ma
2SG	LO	-sa	FA	-sa
3SG	LO	-ta	FA	-ta
1PL	LO	-mo	FA	-mo
2PL	LO	-so	FA	-so
3PL	LO	-to	FA	-to

Traits	LEXÈME 3		LEXÈME 4	
1SG	VO	-mi	ZE	-mi
2SG	VO	-si	ZE	-si
3SG	VO	-ti	ZE	-ti
1PL	VO	-mu	ZE	-mu
2PL	VO	-su	ZE	-su
3PL	VO	-tu	ZE	-tu

FIGURE 3.5 : Deux classes flexionnelles

Le latin est une langue qui fait ainsi usage de classes flexionnelles. Le tableau 3.9 montre deux des classes flexionnelles nominales du latin.

	FILIUS « <i>fil</i> s »		ROSA « <i>rose</i> »	
	SG	PL	SG	PL
NOM	FILI- <i>us</i>	FILI- <i>i</i>	ROS- <i>a</i>	ROS- <i>e</i>
VOC	FILI- <i>e</i>	FILI- <i>i</i>	(ROS- <i>a</i> )	(ROS- <i>e</i> )
ACC	FILI- <i>um</i>	FILI- <i>os</i>	ROS- <i>am</i>	ROS- <i>as</i>
GEN	FILI- <i>i</i>	FILI- <i>orum</i>	ROS- <i>e</i>	ROS- <i>arum</i>
DAT	FILI- <i>o</i>	FILI- <i>is</i>	ROS- <i>e</i>	ROS- <i>is</i>
ABL	FILI- <i>o</i>	FILI- <i>is</i>	ROS- <i>a</i>	ROS- <i>is</i>

TABLE 3.9 : Classes flexionnelles en latin

Les données du latin du tableau 3.9 ne respectent pas le critère de canonicité n° 3 en ce que les cases appartenant à des paradigmes de deux lexèmes donnés mais exprimant les mêmes traits morphosyntaxiques ne comportent pas forcément le même exposant, cf. le datif singulier est marqué par le suffixe *-o* pour FILIUS « *fil*s », mais par le suffixe *-e* pour ROSA « *rose* ».

### 3.5.1.3 Synchrétisme

Mais la manifestation de classes flexionnelles n'est pas le point seul par lequel les données nominales du latin contreviennent à la contrainte n° 3 : la contrainte n° 3 impose qu'à l'intérieur d'un même paradigme, il n'y ait pas deux formes affichant le même exposant. Or, dans le paradigme de *FILIUS*, le DAT.SG et le ABL.SG sont marqués par le même suffixe *-o*, tout comme le GEN.SG, le NOM.PL et le VOC.PL (*-i*) et le DAT.PL et le ABL.PL (*-is*).

Le phénomène consistant à présenter deux cases d'un paradigme formellement identiques est appelé SYNCRÉTISME. Une représentation de ce concept est également donnée dans la figure 3.6 : dans le paradigme du lexème 1, la case de la première personne du singulier est synchrétique avec celle de la deuxième personne du pluriel, et les cases de la troisième personne (singulier et pluriel) sont également synchrétiques entre elles.

Traits	LEXÈME 1	LEXÈME 2
1SG	LO <i>-ma</i>	FA <i>-ma</i>
2SG	LO <i>-sa</i>	FA <i>-sa</i>
3SG	LO <i>-ta</i>	FA <i>-ta</i>
1PL	LO <i>-mo</i>	FA <i>-mo</i>
2PL	LO <i>-ma</i>	FA <i>-so</i>
3PL	LO <i>-ta</i>	FA <i>-to</i>

FIGURE 3.6 : Paradigmes présentant des synchrétismes (lexème 1) et canonique (lexème 2)

### 3.5.1.4 Hétéroclise

L'HÉTÉROCLISE désigne le phénomène où le paradigme d'un lexème se compose de parties d'au moins deux classes flexionnelles distinctes.

Ainsi, dans une langue qui posséderait les classes flexionnelles de notre exemple artificiel définies dans la figure 3.5, un cinquième lexème qui formerait ses première et deuxième personnes selon la classe 1, comme les lexèmes 1 et 2, et sa troisième personne selon la classe 2, comme les lexèmes 3 et 4, serait un exemple d'hétéroclise illustré à la figure 3.7.

Les noms d'animaux slovaque fournissent un exemple d'hétéroclise (Sagot & Walther, 2011 ; Walther & Sagot, 2011). En slovaque, il existe des classes flexionnelles différentes pour les noms animés (cf. *CHLAP* « homme, type » dans le tableau 3.10) et les noms inanimés (cf. *DUB* « chêne »). En revanche, certains noms d'animaux masculins comme *OROL* « aigle » se fléchissent comme des noms masculins animés au singulier, mais peuvent (voire doivent, pour certains lexèmes) être fléchis comme des inanimés au pluriel<sup>36</sup> (Zauner, 1973). Le

36. Ils peuvent néanmoins être fléchis comme des animés dans des cas de personnification explicite. Ce cas

Traits	CLASSE 1	CLASSE 2	LEXÈME 5
1SG	RADICAL - <i>ma</i>	RADICAL - <i>mi</i>	GLU - <i>ma</i>
2SG	RADICAL - <i>sa</i>	RADICAL - <i>si</i>	GLU - <i>sa</i>
3SG	RADICAL - <i>ta</i>	RADICAL - <i>ti</i>	GLU - <i>ti</i>
1PL	RADICAL - <i>mo</i>	RADICAL - <i>mu</i>	GLU - <i>mo</i>
2PL	RADICAL - <i>so</i>	RADICAL - <i>su</i>	GLU - <i>so</i>
3PL	RADICAL - <i>to</i>	RADICAL - <i>tu</i>	GLU - <i>tu</i>

FIGURE 3.7 : Exemple d'hétéroclise

	MASCULIN ANIMÉ CHLAP « <i>homme, type</i> »		MASCULIN INANIMÉ DUB « <i>chêne</i> »		MASCULIN HÉTÉROCLITE OROL « <i>aigle</i> »	
	SG	PL	SG	PL	SG	PL
NOM	<i>chlap</i>	<i>chlap-i</i>	<i>dub</i>	<i>dub-y</i>	<i>orol</i>	<i>orol-y</i>
GEN	<i>chlap-a</i>	<i>chlap-ov</i>	<i>dub-a</i>	<i>dub-ov</i>	<i>orol-a</i>	<i>orol-ov</i>
DAT	<i>chlap-ovi</i>	<i>chlap-om</i>	<i>dub-u</i>	<i>dub-om</i>	<i>orol-ovi</i>	<i>orol-om</i>
ACC	<i>chlap-a</i>	<i>chlap-ov</i>	<i>dub</i>	<i>dub-y</i>	<i>orol-a</i>	<i>orol-y</i>
LOC	<i>chlap-ovi</i>	<i>chlap-och</i>	<i>dub-e</i>	<i>dub-och</i>	<i>orol-ovi</i>	<i>orol-och</i>
INS	<i>chlap-om</i>	<i>chlap-mi</i>	<i>dub-om</i>	<i>dub-mi</i>	<i>orol-om</i>	<i>orol-ami</i>

TABLE 3.10 : Hétéroclise dans les paradigmes des noms d'animaux slovaques

tableau 3.10 indique les paradigme des trois noms CHLAP « *homme, type* », DUB « *chêne* » et OROL « *aigle* »<sup>37</sup>.

### 3.5.2 Non-canonicté du remplissage des cases des paradigmes

Les paramètres 5 et 6 ci-dessus indiquent la manière dont doivent être remplies les cases d'un paradigme canonique : le CRITÈRE DE COMPLÉTUDE indique qu'il ne doit y avoir aucune case vide dans un paradigme canonique, le CRITÈRE D'UNICITÉ spécifie que chacune de cases doit être remplie par exactement une forme. Deux phénomènes non canoniques correspondent au non-respect de ces contraintes, à savoir respectivement la DÉFECTIVITÉ et la SURABONDANCE.

reste néanmoins l'expression volontaire d'un effet de style particulier.

37. CHLAP et DUB se fléchissent de façon régulière : CHLAP appartient à la classe de flexion par défaut des noms masculins animés qui se terminent par une consonne, tandis que DUB appartient à classe standard pour les inanimés masculins se terminant par ce qu'on désigne des consonnes *dures* ou *neutres* dans la tradition linguistique slave.

## 3.5.2.1 Défectivité

La DÉFECTIVITÉ désigne les cas où un lexème possède dans son paradigme des cases vides (ou manquantes), cf. l'illustration dans la figure 3.8 (défectivité à la deuxième personne du pluriel pour le lexème 1).

Traits	LEXÈME 1	LEXÈME 2
1SG	LO <i>-ma</i>	FA <i>-ma</i>
2SG	LO <i>-sa</i>	FA <i>-sa</i>
3SG	LO <i>-ta</i>	FA <i>-ta</i>
1PL	LO <i>-mo</i>	FA <i>-mo</i>
2PL		FA <i>-so</i>
3PL	LO <i>-to</i>	FA <i>-to</i>

FIGURE 3.8 : Défectivité à la deuxième personne du pluriel (lexème 1)

Les langues possèdent ainsi des lexèmes pour lesquels des formes pourtant attendues n'existent pas. Quand bien même un contexte appellerait l'emploi d'une de ces formes manquantes, les locuteurs sont dans l'incapacité de les produire.

Un exemple classique de défectivité en français est le verbe GÉSIR qui, dans son paradigme, n'affiche que 12 formes sur les 44 formes (non composées) habituellement affichées par les verbes. Un autre exemple, plus extrême encore, est le verbe CHALOIR qui ne possède que deux formes, à savoir l'infinitif *chaloir* et la troisième personne du singulier (3SG.PRS.IND) *chaut*.

Le problème principal dans la définition de la défectivité réside dans la définition de ce qu'est une forme inexistante. En effet, dès lors qu'une forme est indisponible pour les locuteurs, ils ont pour habitude d'avoir recours à des STRATÉGIES DE RÉPARATION. Ainsi la question se pose-t-elle si le recours à des phrases périphrastiques doit être considéré comme une réparation interne ou externe au paradigme.

Les deux exemples ci-dessus sont des cas de défectivité isolés au sein de l'ensemble des verbes français. Mais la défectivité peut également affecter des classes entières de lexèmes comme dans le cas des noms communément désignés comme PLURALIA TANTUM. Ces noms n'existent que sous leur forme plurielle. Nous en trouvons des exemples dans diverses langues, cf. en français (VIVRES), en anglais ( cf. TROUSERS « *pantalon* »), ou encore en slovaque (cf. VIANOCE « *Noël* »). Il est à noter que pour les langues comme le slovaque dans lesquelles les paradigmes nominaux affichent une alternance casuelle, la défectivité en nombre affecte chacun des cas. Le paradigme de VIANOCE est donc diminué de moitié du fait de l'absence des formes du singulier (cf. les six cases du paradigme de VIANOCE illustré

par le tableau 3.11 *vs.* douze cases des paradigmes du tableau 3.10 pour les noms non défectifs CHLAP « *homme, type* », DUB « *chêne* » et OROL « *aigle* »).

	SG	PL
NOM		<i>Vianoc-e</i>
GEN		<i>Vianoc</i>
DAT		<i>Vianoc-iam</i>
ACC		<i>Vianoc-e</i>
LOC		<i>Vianoc-iach</i>
INS		<i>Vianoc-ami</i>

TABLE 3.11 : Exemple de paradigme défectif slovaque : VIANOCE

### 3.5.2.2 Surabondance

La contrepartie de la défectivité est la notion de SURABONDANCE. On parle de surabondance à chaque fois qu'une case d'un paradigme donné contient plus d'une forme (cf. l'illustration de la figure 3.9 pour les formes du lexème 1).

Traits	LEXÈME 1		LEXÈME 2	
1SG	LO	<i>-ma</i>	FA	<i>-ma</i>
	LO	<i>-man</i>		
2SG	LO	<i>-sa</i>	FA	<i>-sa</i>
	LO	<i>-san</i>		
3SG	LO	<i>-ta</i>	FA	<i>-ta</i>
	LO	<i>-tan</i>		
1PL	LO	<i>-mo</i>	FA	<i>-mo</i>
	LO	<i>-mon</i>		
2PL	LO	<i>-so</i>	FA	<i>-so</i>
	LO	<i>-son</i>		
3PL	LO	<i>-to</i>	FA	<i>-to</i>
	LO	<i>-ton</i>		

FIGURE 3.9 : Paradigme surabondant (lexème 1)

La notion de surabondance contrevient au critère n° 6 d'unicité de forme par case. Elle a été introduite par Thornton (2010, 2011, 2012) pour l'italien. La surabondance concerne canoniquement les cas où deux COMPAGNONS DE CELLULE (*cell mates*) d'une même case sont en compétition pour la réalisation des traits associés, sans qu'aucun trait particulier ne permette de choisir l'un plutôt que l'autre. Le tableau 3.12 montre des exemples de surabondance présentés par Thornton pour l'italien.

LEXÈME	TRAITS MORPHOSYNTAXIQUES	COMPAGNON 1	COMPAGNON 2
« languir »	3PL.PRS.SBJV	<i>languano</i>	<i>languiscano</i>
« posséder »	3PL.PRS.SBJV	<i>possiedano</i>	<i>posseggano</i>
« posséder »	3SG.PRS.SBJV	<i>possieda</i>	<i>possegga</i>
« posséder »	1SG.PRS.SBJV	<i>possiedo</i>	<i>posseggo</i>

TABLE 3.12 : Surabondance en italien d'après Thornton (2010)

Un exemple en français est le verbe ASSEOIR qui possède deux formes distinctes dans la majorité de ses cases<sup>38</sup> (cf. le tableau 3.13).

IND.PRS	SG	PL
1	<i>assis</i>	<i>assoyons</i>
	<i>assieds</i>	<i>asseyons</i>
2	<i>assis</i>	<i>asseyez</i>
	<i>assieds</i>	<i>asseyez</i>
3	<i>assoit</i>	<i>assoient</i>
	<i>assied</i>	<i>asseyent</i>

TABLE 3.13 : Surabondance en français pour le verbe ASSEOIR

De même, tous les verbes français en *-ayer* affichent une surabondance systématique (cf. le tableau 3.14). En effet, dans certaines cases, ces verbes affichent deux radicaux concurrents, à savoir l'un en *-ay-* et l'autre en *-ai-*. Ils ont donc à chaque fois deux formes concurrentes, équivalentes morphologiquement (même s'il peut y avoir dans ces cas des différences d'ordre sémantique, pragmatique, sociolectales ou autres).

### 3.5.3 Non-canonicté dans l'expression des traits morphosyntaxiques

Les deux derniers paramètres de canonicité que nous avons évoqués (n° 7 et 8) concernent le mode de RÉALISATION DES TRAITS MORPHOSYNTAXIQUES. Comme l'indiquent Brown & Hippisley (2012), la flexion canonique a pour but de construire les formes réalisant des traits morphosyntaxiques imposés par la syntaxe. Cette réalisation se fait par ailleurs de façon régulière : un paquets de traits morphosyntaxiques donné sera toujours réalisé de la même manière pour une catégorie morphosyntaxique donnée.

38. Voir Bonami & Boyé (2010a) pour plus de détails à ce sujet.



	SG	PL
1	<i>balaye</i> <i>balaie</i>	<i>balayons</i>
2	<i>balayes</i> <i>balaies</i>	<i>balayez</i>
3	<i>balaye</i> <i>balaie</i>	<i>balayent</i> <i>balaient</i>

TABLE 3.14 : Surabondance à l'indicatif présent en français pour le verbe BALAYER

Canoniquement, chaque paradigme d'un lexème d'une même catégorie morphosyntaxique comporte donc le même nombre de cases dont les formes réalisent le même inventaire de traits morphosyntaxiques. La non-canonicté peut donc s'exprimer par la réalisation d'un nombre inférieur ou supérieur de traits ou par une inadéquation entre un exposant affiché et les traits réalisés.

### 3.5.3.1 Surdifférenciation de valeurs et de traits

Certains lexèmes affichent en revanche dans leur paradigme des formes exprimant des distinctions qui ne sont pas attendues compte tenu des traits habituellement exprimés par la catégorie morphosyntaxique à laquelle ils appartiennent. Le paradigme résultant présente alors une quantité surnuméraire de cases distinctes. Le phénomène non canonique en question est appelé SURDIFFÉRENCIATION. On parle de surdifférenciation quand un lexème donné distingue plus de traits ou plus de valeurs pour un trait donné que ce qui serait attendu compte tenu des traits ou valeurs distingués par les lexèmes de la classe morphosyntaxique à laquelle il appartient.

Un cas de sur différenciation peut être illustré par le comportement des noms masculins de la classe flexionnelle de FILIUS « *flis* » illustrée au tableau 3.9 : la plupart des noms latins ne distinguent qu'entre cinq cas (NOM, ACC, GEN, DAT, ABL). Seuls les noms de la classe de FILIUS distinguent de surcroît entre le nominatif (en *-us*) et le vocatif (en *-e*). Cette classe de nom est cependant assez importante. En revanche, on observe ce même phénomène de façon plus extrême encore pour le slovaque où seuls une douzaine de noms<sup>39</sup> affichent un vocatif en plus des six cas habituels (NOM, GEN, DAT, ACC, LOC, INS) (Zauner, 1973).

39. SYN « *flis* » (vocatif *synku*), BRAT « frère » (vocatif *brat(k)u*), CHLAPEC « garçon » (vocatif *chlapče*), ŠVAGOR « beau-frère » (vocatif *švagre*), KMOTOR « parrain » (vocatif *kmotre*), CHLAP « homme » (vocatif *chlape*), PRIATEL « ami » (vocatif *priatel'u*), PÁN « Monsieur » (vocatif *pane*), MAJSTER « maître » (vocatif *majtre*), BOH « dieu » (vocatif *bože*), ČLOVEK « individu » (vocatif *človece*).

	NEUTRE EN –ET DUGME « bouton »		MASCULIN JELEN « cerf »		FÉMININ EN –A ŽENA « femme »	
	SG	PL	SG	PL	SG	PL
NOM	<i>dugme</i>	<i>dugm-et-a</i>	<i>jelen</i>	<i>jelen-i</i>	<i>žen-a</i>	<i>žen-e</i>
ACC	<i>dugme</i>	<i>dugm-et-a</i>	<i>jelen-a</i>	<i>jelen-e</i>	<i>žen-u</i>	<i>žen-e</i>
GEN	<i>dugm-et-a</i>	<i>dugm-et-a</i>	<i>jelen-a</i>	<i>jelen-a</i>	<i>žen-e</i>	<i>žen-a</i>
DAT	<i>dugm-et-u</i>	<i>dugm-et-ima</i>	<i>jelen-u</i>	<i>jelen-ima</i>	<i>žen-i</i>	<i>žen-ama</i>
LOC	<i>dugm-et-u</i>	<i>dugm-et-ima</i>	<i>jelen-u</i>	<i>jelen-ima</i>	<i>žen-i</i>	<i>žen-ama</i>
INS	<i>dugm-et-om</i>	<i>dugm-et-ima</i>	<i>jelen-om</i>	<i>jelen-ima</i>	<i>žen-om</i>	<i>žen-ama</i>

TABLE 3.15 : Déclinaison régulière de noms du serbo-croate

### 3.5.3.2 Décalage (« déponence »)

Le dernier phénomène non canonique que nous décrivons ici est le phénomène de DÉCALAGE (entre forme et fonction).

Un exemple de décalage peut être trouvé dans certains paradigmes de noms serbo-croates. Ils se fléchissent en nombre en suivant un certain nombre de classes flexionnelles distinctes. Parmi ces classes, nous nous intéressons ici plus précisément à celles des féminins à RADICAL EN –A comme ŽENA « femme » et masculins comme JELEN « cerf » respectivement, ainsi que celle des neutres à RADICAL EN –ET comme DUGME « bouton », présentées dans le tableau 3.15 (Baerman, 2006).

Les noms DETE « enfant » et BRAT « frère » sont des noms qui se fléchissent respectivement selon les patrons du singulier des noms à RADICAL EN –ET et des noms à RADICAL EN –A, cf. le tableau 3.16. Cependant, ils font partie de ces noms du serbo-croate qui emploient des formes du singulier d’une autre classe pour exprimer leur pluriel, à savoir en empruntant les désinences du singulier des noms féminins en –A. Le fait d’employer le singulier (d’une autre classe ou non) pour exprimer le pluriel constitue un décalage entre forme et fonction <sup>40</sup> (Baerman, 2007b).

Ainsi, les formes d’un paradigme présentent parfois un décalage par rapport à la fonction morphosyntaxique réalisée. Baerman (2007b) choisit de donner à ce décalage entre forme et fonction le nom de DÉPONENCE, d’après les verbes déponents latins souvent décrits comme exprimant la voix active à l’aide de formes passives. Pourtant, comme nous l’avons montré au chapitre 1, le terme de « déponent » porte ici à confusion. Dans le cas particulier des « déponents latins », il n’y a justement pas de décalage comme celui illustré par les nom du serbo-croate ci-dessus, la voix n’étant pas un trait définitoire des paradigmes latins <sup>41</sup>.

40. « A mismatch between form and function. »

41. Nous montrons au chapitre 1 que la passivation constitue une opération dérivationnelle en latin.

	NEUTRE EN –ET/ RADICAL EN –A DETE « enfant »		MASCULIN / RADICAL EN –A BRAT « frère »	
	SG	PL	SG	PL
NOM	<i>dete</i>	<i>dec-a</i>	<i>brat</i>	<i>brać-a</i>
ACC	<i>dete</i>	<i>dec-u</i>	<i>brat-a</i>	<i>brać-u</i>
GEN	<i>det-et-a</i>	<i>dec-e</i>	<i>brat-a</i>	<i>brać-e</i>
DAT	<i>det-et-u</i>	<i>dec-i</i>	<i>brat-u</i>	<i>brać-i</i>
LOC	<i>det-et-u</i>	<i>dec-i</i>	<i>brat-u</i>	<i>brać-i</i>
INS	<i>det-et-om</i>	<i>dec-om</i>	<i>brat-om</i>	<i>brać-om</i>

TABLE 3.16 : Décalage dans deux paradigmes de noms du serbo-croate

Nous proposons donc d'employer le terme de *décalage* comme dénomination de ce type de phénomène non canonique et de garder le terme de « déponence » comme dénomination des verbes hétéroclites latins du type du verbe LOQUOR « parler »<sup>42</sup>.

### 3.5.4 Huit phénomènes non canoniques : résumé

La figure 3.10 représente les huit phénomènes de non-canonlicité que nous avons exposés dans cette section. Chaque phénomène non canonique y est représenté à l'extrémité d'un axe s'éloignant du centre représentant le canon. Ainsi, le fait d'afficher les caractéristiques définitoires de chacun de ces phénomènes (définies plus bas) de façon plus ou moins importante correspond à une distance plus ou moins grande par rapport au centre du cercle. Nous avons étiqueté chacune des branches par une lettre (F, C, T) correspondant au type d'irrégularités observables pour chacun des phénomènes : F pour des irrégularités portant sur la réalisation des formes du paradigme, C pour les irrégularités dans le remplissage des cases du paradigme et T pour les irrégularité concernant la réalisation des traits morphosyntaxiques. Le tableau 3.17 page 186 résume les caractéristiques de chacun des phénomènes non canoniques indiqués dans la figure 3.10 page 185.

Le modèle formel que nous présentons au chapitre 4 a été conçu pour rendre compte des huit déviations par rapport à la norme canonique que nous avons évoquées dans cette section : la supplétion de radical et de forme, le syncrétisme, l'hétéroclise, la défektivité, la surabondance, la surdifférenciation et le décalage morphosyntaxique. Pour chacun de ces phénomènes, nous montrerons que la modélisation d'un phénomène donné à

42. Notons toutefois que le *Surrey Morphology Group* a constitué une base de donnée étendue de phénomènes de décalage. Cette base de données constitue une excellente présentation de divers phénomènes de décalage observables dans les langues du monde. Elle est accessible en ligne à l'adresse <http://www.smg.surrey.ac.uk/deponency> (valable au 11 juin 2013).

FIGURE 3.10 : Axes de non-canonicit 

l'int rieur d'un syst me mod lis  d pend des comportements par d faut qu'affiche une langue donn e.

### 3.6 Prototypes, formes naturelles et canon en typologie

Nous avons indiqu  au d but de ce chapitre que les deux objectifs majeurs de la typologie sont 1) de recenser les diff rences observables entre les langues et 2) de faire l'inventaire des invariants linguistiques. Ces deux objectifs peuvent  tre r alis s   l'aide des diff rents moyens d crits dans ce chapitre.

Dans la section 3.1 nous avons r sum  comment,   partir d' tudes qualitatives sur un maximum de langues donn es, peut  tre mise en place une architecture de concepts th oriques exploitables en typologie.

Une fois les concepts d finis, une  tude quantitative de la r alisation de ces concepts permet d'identifier des tendances g n rales et d' tablir un inventaire d'universaux permettant de d duire ce que serait, selon l'observation des langues du monde, un syst me

	PHÉNOMÈNE	DESCRIPTION
F1	SUPLÉTION DE RADICAL	<i>Le paradigme affiche des alternances de radical. La définition stricte de la supplétion veut que cette alternance concerne des formes non prévisibles.</i>
F2	SUPLÉTION DE FORME	<i>Le paradigme comporte des formes qui remplacent à la fois le radical et les exposants attendus pour une case donnée.</i>
F3	SYNCRÉTISME	<i>Le paradigme comporte deux formes identiques.</i>
C1	HÉTÉROCLISE	<i>Les formes du paradigme sont construites à partir d'au moins deux classes flexionnelles distinctes.</i>
C2	DÉFECTIVITÉ	<i>Le paradigme comporte des cases vides.</i>
C3	SURABONDANCE	<i>Le paradigme comporte des cases avec plus d'une forme.</i>
T1	SURDIFFÉRENCIATION	<i>Le paradigme comporte un surplus de cases dû à la distinction de traits ou des valeurs de traits supplémentaires.</i>
T2	DÉCALAGE	<i>Un ou plus d'un paquet de traits morphosyntaxiques est exprimé par une forme attendue pour un autre paquet de traits morphosyntaxiques.</i>

TABLE 3.17 : Récapitulatif des huit types de phénomènes non canoniques

prototypique (flèches bleues).

Les tendances générales entrent également en compte pour la définition des paramètres ordonnés définissant un système naturel (flèches vertes). Ces paramètres sont en outre définis par des contraintes extra-linguistiques (sémiotiques, phonologiques, ...) sur la réalisation des concepts théoriques. L'ordonnement de ces paramètres est ainsi le fruit de ces contraintes. Mais le système le plus naturel est également conditionné par l'importance quantitative d'un mode de réalisation d'un concept pour une langue donnée

(cf. la différence entre naturalité globale et locale, section 3.3.3).

Enfin, le standard canonique est l'application directe de la réalisation la plus régulière possible des concepts théoriques définis, telle que cette réalisation permette une absence totale d'ambiguïté (flèches rouges).

Le schéma 3.1 résumait la mise en place de ces différents standards. Dans le schéma 3.11 nous les mettons en rapport avec les deux objectifs de la typologie. D'une part les universaux identifiés par les études quantitatives permettent-elles de lister les invariants interlangues (absolus ou tendanciels) observés. D'autre part, les trois types de standards construits comme indiqués dans le schéma 3.1 peuvent servir de repères auxquels les différentes langues attestées peuvent être comparées.

Chacun de ces standards a par ailleurs une utilité propre. Les différences par rapport au système prototypique marquent combien inhabituelle la structure d'une langue est par rapport aux tendances connues jusqu'ici. Le rapport particulier qu'une langue entretient avec la norme naturelle permet de montrer comment les contraintes extralinguistiques peuvent interagir de façon différentes pour créer des systèmes linguistiques formellement différents. Dans une démarche fonctionnelle, le standard naturel permet également de mesurer l'adéquation d'une forme avec les fonctions qu'elle est censée remplir dans la langue. Enfin, une déviation par rapport au système canonique marque et le degré d'irrégularité d'une langue et le domaine dans lequel cette irrégularité se manifeste.

Contrairement au prototype, soumis aux tendances observables à travers les langues, le canon, indice de régularité maximale et d'ambiguïté minimale, n'est pas susceptible de varier en fonction de découvertes de nouvelles langues ou de nouveaux phénomènes observables. De ce point de vue il fait ainsi preuve d'une plus grande stabilité. Par rapport au critère de naturalité, soumise aux interactions avec des modules extérieurs au système étudié, il constitue un standard plus autonome et moins dépendant du poids accordé à chacun de ces modules par rapport au système et les uns par rapport aux autres.

La notion de canonicité est ainsi particulièrement adaptée à une approche lexicaliste forte assumant pleinement l'autonomie de la morphologie par rapport aux autres modules de la grammaire, c'est-à-dire comme un système complexe à décrire indépendamment. Elle se prête également particulièrement à la conception d'un modèle formel pour la typologie morphologique dont les primitives doivent nécessairement être fixées. Enfin, grâce à son caractère indépendant, elle permet aisément une évaluation interne des caractéristiques proprement morphologiques des systèmes étudiés.

Toutefois, les approches canoniques de la morphologie, quoique justement conçues comme cadre pour des mesures de degré de canonicité, ne proposent pas de mesures

FIGURE 3.11 : Approches en typologie

explicités de la canonicité. En effet, il est nécessaire, avant de mesurer, de pouvoir représenter, isoler et définir formellement les paramètres à mesurer. En d'autres termes, un modèle formel de la canonicité est nécessaire pour effectuer des mesures de canonicité.

Au chapitre 4, nous introduisons un modèle formel de la morphologie flexionnelle conçu pour rendre compte des déviations d'un système flexionnel par rapport au canon. Ce modèle s'accompagne également de mesures permettant de quantifier le degré de déviation par rapport au canon, et ce phénomène non canonique par phénomène et de dresser un portrait quantifié du type de non-canonicité dont fait preuve le système flexionnel d'une langue donnée.





## Deuxième partie

Formaliser une approche typologique  
de la morphologie flexionnelle



## PARSLI, un modèle formel de la morphologie flexionnelle

L'approche canonique de la typologie morphologique décrite au chapitre 3 permet d'évaluer qualitativement les déviations par rapport à un canon défini. Mais quand bien même le concept de canonicité est essentiellement une théorisation de la déviation de la régularité formelle, elle ne s'accompagne à ce jour d'aucun modèle formel permettant de formaliser en tant que telles les déviations par rapport à ce canon, de les évaluer finement et, à termes, de les quantifier. C'est entre autres ce manque que nous tentons de pallier en proposant dans ce chapitre un nouveau modèle formel de la morphologie flexionnelle.

Le modèle PARSLI que nous introduisons ici a été spécifiquement conçu pour rendre compte des primitives qui sous-tendent l'approche canonique de la morphologie. Il propose une représentation explicite de la structure du paradigme de chaque entrée lexicale d'une langue donnée ainsi que de sa déviation par rapport au canon. La structure du paradigme d'un lexème  $y$  est conçue comme la conséquence des éléments constitutifs de la définition de son entrée lexicale spécifique. L'étendue de chaque phénomène non canonique dans un paradigme donné peut ainsi être identifiée à partir des informations associées à son entrée lexicale.

Par la formalisation des entrées lexicales, PARSLI permet de caractériser les propriétés flexionnelles non canoniques de chaque lexème. À partir de cette formalisation que propose PARSLI, nous pouvons de surcroît définir des MESURES DE CANONICITÉ permettant de rendre compte, pour une description donnée, du degré de non-canonycité d'un paradigme particulier ou d'un système flexionnel dans son ensemble, et ce explicitement phénomène par phénomène. Il permet ainsi d'étendre la puissance de l'approche canonique de la morphologie flexionnelle en lui offrant la possibilité de non

---

seulement qualifier les différences d'un système flexionnel donné par rapport au canon, mais également de les quantifier.

Mais au-delà de la formalisation de la flexion canonique,  $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}\mathcal{I}$  constitue UN MODÈLE DE MORPHOLOGIE FLEXIONNELLE indépendant, et en particulier un modèle du rapport entre lexicque et structures paradigmatiques, comme l'indique son nom acronymique  $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}\mathcal{I}$ , signifiant «  $\mathcal{P}\mathcal{A}\mathcal{R}$ adigm  $\mathcal{S}$ hape and  $\mathcal{L}$ exicon  $\mathcal{I}$ nterface »<sup>1</sup>.

L'objectif de  $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}\mathcal{I}$  est de rendre compte explicitement du rapport entre la structuration des entrées lexicales dans le LEXIQUE FLEXIONNEL d'une part et l'organisation des opérations de la GRAMMAIRE FLEXIONNELLE, notamment en termes de classes flexionnelles, de l'autre. Il comporte ainsi une véritable théorie du lexique qui définit la structure de la composante flexionnelle d'une entrée lexicale, ainsi qu'une théorie de la morphologie flexionnelle qui comporte notamment une théorie des classes flexionnelles et de l'allomorphie radicale.

$\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}\mathcal{I}$  est un modèle INFÉRENTIEL-RÉALISATIONNEL au sens de Stump (2001) s'inscrivant dans une approche de la morphologie de type MOT ET PARADIGME. Il s'insère dans une VISION MODULAIRE ARTICULÉE DE LA GRAMMAIRE<sup>2</sup>, laquelle suppose à la fois une AUTONOMIE réelle de la morphologie flexionnelle par rapport aux autres composantes de la grammaire, notamment de la syntaxe, et une interaction complexe avec ces dernières, notamment avec un lexique structuré.

$\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}\mathcal{I}$  permet de rendre compte de la structure complète d'un système flexionnel depuis son ancrage dans le lexique jusque dans la grammaire sous-jacente à la structure des paradigmes. Il est de ce point de vue à comparer avec d'autres modèles de la morphologie flexionnelle comme PFM (Stump, 2001, 2006) et le volet flexionnel de *Network Morphology* (Brown & Hippisley, 2012).

Nous présentons ici en un premier temps, à la section 4.1, les caractéristiques globales de  $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}\mathcal{I}$  et notamment les points par lesquels il se distingue d'autres modèles existants avec lesquels il partage un certain nombre de présupposés théoriques et d'outils analytiques. Nous présenterons ensuite les primitives du modèle et leur rapport avec les notions de canonicité (section 4.2). À la section 4.4, nous introduisons les MESURES DE CANONICITÉ définies à partir des notions disponibles dans  $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}\mathcal{I}$ . Une application de ces mesures à une description complète sera proposée au chapitre 7.

---

1. Une première version de  $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}\mathcal{I}$  a été publiée dans (Walther, 2011b) et figure également plus schématiquement dans (Sagot & Walther, 2011) et (Walther & Sagot, 2011) en tant que modèle sous-jacent aux expériences décrites dans ces deux articles. Le modèle a néanmoins évolué depuis. La version que nous proposons dans ce chapitre se distingue par conséquent en un certain nombre de points de la version préliminaire décrite dans (Walther, 2011b). La notation a notamment été simplifiée.

2. Cf. section 2.2, chapitre 2.

## 4.1 Description des propriétés spécifiques de $\mathcal{PARSLI}$

### 4.1.1 Zones flexionnelles

Un des points principaux par lesquels  $\mathcal{PARSLI}$  se distingue d'autres modèles existants, et notamment de PFM et *Network Morphology*, est la conception des CLASSES FLEXIONNELLES qu'il encode : contrairement à ces autres modèles,  $\mathcal{PARSLI}$  n'associe pas une entrée lexicale à une classe flexionnelle, mais à un ensemble de ZONES FLEXIONNELLES qui définissent la réalisation de sous-parties de son paradigme. Dans  $\mathcal{PARSLI}$ , les classes flexionnelles sont définies à partir des zones flexionnelles et non l'inverse.

	SG	PL
1	<i>balaye</i> <i>balaie</i>	<i>balayons</i>
2	<i>balayes</i> <i>balaies</i>	<i>balayez</i>
3	<i>balaye</i> <i>balaie</i>	<i>balayent</i> <i>balaient</i>

TABLE 4.1 : Surabondance à l'indicatif présent en français pour le verbe BALAYER

La notion de zone flexionnelle permet d'exprimer des généralisations que ne captent pas les autres approches de type mot et paradigme. Prenons par exemple la flexion verbale du français<sup>3</sup>. Parmi les verbes du premier groupe, les verbes en *-ayer* manifestent une SURABONDANCE RÉGULIÈRE, illustrée pour l'indicatif présent du verbe BALAYER au tableau 4.1. Dans (Bonami & Boyé, 2003) et (Bonami & Boyé, 2010a, pp. 5-6), les auteurs considèrent ces verbes comme *polyparadigmatiques*<sup>4</sup>. Chacun des verbes en *-ayer* est associé à deux jeux de radicaux, l'un ne comportant qu'un radical unique en *ej*, l'autre comportant

3. La flexion verbale du français est traditionnellement décrite comme comportant une classe flexionnelle régulière et productive, celle des verbes du premier groupe (verbes en *-er*), une classe flexionnelle irrégulière (verbes du troisième groupe), et la classe flexionnelle des verbes du deuxième groupe (verbes comme FINIR), parfois considérée comme régulière et parfois comme irrégulière. Les analyses diffèrent sur la régularité de cette dernière classe (Boyé, 2000 ; Kilani-Schoch & Dressler, 2005 ; Bonami et al., 2008), ce qui constitue un premier point possible de divergence entre les différentes analyses possibles de la flexion verbale du français.

4. « La plupart des lexèmes ont un paradigme stable, chaque case du paradigme étant clairement remplie par une forme et une seule, mais il y a des exceptions. Un cas particulièrement parlant est celui du verbe *asseoir* en français. [...] Le cas d'ASSEOIR est particulièrement net, mais il n'est pas isolé. Le verbe HAÏR connaît les présents singuliers *je hais* et *je haïs* ; le verbe PAYER dispose de *je paie* et *je paye* ; etc. [...] Il est donc rare, mais pas impossible, qu'un lexème dispose de plusieurs paradigmes en concurrence. Cette situation constitue une exception claire à la fonctionnalité de la flexion : pour un lexème et un paquet de traits morphosyntaxiques de traits donnés, il y a plus d'une forme possible. (Bonami & Boyé, 2010a, pp. 5-6). »

deux radicaux (en  $\epsilon_j$  et en  $\epsilon$ ). Chacun de ces jeux de radicaux est fléchi selon les règles de flexion des verbes du premier groupe en français. Le résultat est la génération des deux paradigmes complets pour un même verbe. Dans ces paradigmes, certaines formes (construites sur un radical en  $\epsilon_j$ ) sont par ailleurs identiques. Par exemple, la première personne du pluriel de l'indicatif *balayons* ( $\text{balej}\tilde{\text{o}}$ ) est obtenue à la fois par la combinaison du radical *balej* du premier jeu de radicaux avec le suffixe  $\tilde{\text{o}}$  et par la combinaison de *balej* du second jeu de radicaux avec  $\tilde{\text{o}}$ . Les formes *balaie* et *balaye* sont en revanche construites par l'emploi de deux radicaux différents : *bale* du premier jeu de radicaux et *balej* du second jeu, chacun associé à une absence de marquage.

BALAYER		
	BALAI-	BALAY-
1SG	<i>balaie</i>	<i>balaye</i>
2SG	<i>balaies</i>	<i>balayes</i>
3SG	<i>balaie</i>	<i>balaye</i>
1PL	<i>balayons</i>	<i>balayons</i>
2PL	<i>balayez</i>	<i>balayez</i>
3PL	<i>balaient</i>	<i>balayent</i>

TABLE 4.2 : Représentation polyparadigmatique des verbes en *-ayer* (exemple de l'indicatif présent de BALAYER)

Cette solution, illustrée au tableau 4.2, n'est pas pleinement satisfaisante dans la mesure où les deux paradigmes ainsi supposés partagent ainsi les mêmes formes pour la moitié des cases de leurs paradigmes <sup>5</sup>.

Une autre façon de représenter cette situation est de définir deux radicaux, l'un en *-ay-* ou  $\epsilon_j$  et l'autre en *-ai-* ou  $\epsilon$  <sup>6</sup>, et de partitionner la classe flexionnelle des verbes du premier groupe (que nous noterons  $Z_{v1}^{exp}$ ) en deux zones flexionnelles : l'une, notée  $z_{v1,1}^{exp}$ , qui est utilisée seulement avec le radical en *-ay-*, et l'autre, notée  $z_{v1,2}^{exp}$ , qui est utilisée par les deux radicaux. Ainsi, nous obtenons trois combinaisons de zones possibles :  $z_{v1,1}^{exp}$  et le radical en *-ay-*,  $z_{v1,2}^{exp}$  et le radical en *-ai-* et  $z_{v1,2}^{exp}$  et le radical en *-ay-*. Dans  $\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}$ , ces trois combinaisons entrent dans la constitution de ce que nous qualifions de sous-

5. Dans l'inventaire de formes syncrétiques indiquées dans le Bescherelle (Arrivé, 1997) par exemple, la proportion de formes (synthétiques) dédoublées est de 21 cases sur 48, soit environ 44% de chacun des paradigmes.

6. Les données de notre analyse sont des données du français écrit utilisées pour les expériences dans (Sagot & Walther, 2011) et (Walther & Sagot, 2011). La même analyse pourrait être faite sur l'oral. Pour ces données, cela ne changerait rien au découpage des radicaux et aurait pour seule conséquence une simplification du système désinentiel.

SCHÈMES, les différents sous-schèmes appartenant à un SCHÈME complet et spécifique à une classe d'entrées lexicales, en l'occurrence à celui des verbes en *-ayer*. Les combinaisons de ces zones avec leur radical permettent de rendre compte du fait que certaines cases des paradigmes des verbes en *-ayer* sont surabondantes, mais certaines seulement. Une telle modélisation a l'avantage de représenter de façon explicite la surabondance d'un verbe par le double emploi de certaines règles de réalisation appartenant à une sous-partie de la classe flexionnelle sans que les formes non surabondantes ne soient générées deux fois. La représentation est ainsi plus compacte et représente plus explicitement la réalité d'un paradigme unique des formes d'un lexème particulier dans lequel certaines cases seulement présentent la propriété non canonique d'être surabondantes. Ce cas de figure est illustré par les figures 4.1 et 4.2 pour les formes de l'indicatif présent de BALAYER.

Classe $Z_{v1}^{exp}$		
1SG	-e	$z_1^{exp}$
2SG	-es	
3SG	-e	$z_2^{exp}$
1PL	-ons	
2PL	-ez	
3PL	-ent	

FIGURE 4.1 : Zones flexionnelles du présent de l'indicatif des verbes du premier groupe en français, c'est-à-dire de la classe flexionnelle  $Z_{v1}^{exp}$

	$balay + z_{v1,1}^{exp}$	$balai + z_{v1,2}^{exp}$	$balay + z_{v1,2}^{exp}$
1SG		<i>balaie</i>	<i>balaye</i>
2SG		<i>balaies</i>	<i>balayes</i>
3SG		<i>balaie</i>	<i>balaye</i>
1PL	<i>balayons</i>		
2PL	<i>balayez</i>		
3PL		<i>balaient</i>	<i>balayent</i>

FIGURE 4.2 : Sous-schèmes (simplifiés) pour le BALAYER

Les radicaux définis à l'intérieur des schèmes peuvent être comparés aux THÈMES regroupés par ESPACES THÉMATIQUES chez Bonami & Boyé (2003) et aux CLASSES DE PARTITIONS<sup>7</sup> de Pirrelli & Battista (2000). Nous les appelons ZONES DE RADICAUX en  $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}\mathcal{I}$ . Elles définissent les champs des règles de sélection de radicaux au sens de PFM (Stump, 2001, 2006). En PFM, ces champs conditionnent chacun une règle de formation de radicaux appliquée à un lexème donné pour l'espace sélectionné.

7. PARTITION CLASSES en anglais.



Les zones flexionnelles sont une généralisation des zones de radicaux au système des exposants. Elles sélectionnent, pour un certain nombre de paquets de traits morphosyntaxiques (PTM) à réaliser par des formes d'un lexème donné, un ensemble de règles de réalisation réalisant ces formes. À la différence d'une zone de radicaux qui ne comporte qu'une règle de formation de radical unique, une zone flexionnelle peut comporter plusieurs règles de réalisation dont chacune participe à la réalisation de la forme d'un lexème pour un PTM<sup>8</sup> donné.

En  $\mathcal{PARSL}$ , les CLASSES FLEXIONNELLES sont les associations de zones les plus souvent sélectionnées conjointement pour la réalisation du paradigme d'un même lexème. De ce point de vue, les classes flexionnelles ne sont pas des primitives d'une description morphologique mais sont dérivées des associations les plus communes de zones flexionnelles selon la structure des paradigmes des lexèmes de la langue.

De même, nous définissons des CLASSES DE RADICAUX comme les associations majoritaires de zones de radicaux.

#### 4.1.2 Représentation de l'hétéroclise

Les zones flexionnelles dans  $\mathcal{PARSL}$  permettent également de rendre compte simplement de l'HÉTÉROCLISE. Rappelons que l'hétéroclise désigne le phénomène où le paradigme d'un lexème se compose de parties d'au moins deux classes flexionnelles distinctes. Dans une modélisation  $\mathcal{PARSL}$ , cela se traduit par l'association d'un lexème donné avec des zones flexionnelles qui ne sont habituellement pas associées à un même lexème et ne constituent donc pas une classe flexionnelle au sens défini plus haut. Les classes flexionnelles sont dérivées des associations de zones flexionnelles observables pour une proportion significative de paradigmes. Aux chapitres 6 et 7, nous reviendrons plus en détail sur les critères (notamment quantitatifs) permettant de décider si un certain ensemble de zones forment ou non entre elles une classe flexionnelle.

Nous avons donné un exemple d'hétéroclise pour certains noms d'animaux en slovaque au chapitre 3. Le tableau 4.3 en rappelle les données pertinentes.

Pour rendre compte des données du tableau 4.3 en  $\mathcal{PARSL}$ , il suffit de considérer que nous sommes face à deux classes flexionnelles distinctes :  $Z_{\text{anim}}^{\text{exp}}$  comportant deux zones flexionnelles par défaut associées aux noms animés slovaques et  $Z_{\text{inan}}^{\text{exp}}$  comportant également deux zones flexionnelles qui sont par défaut associées aux noms inanimés. Les zones flexionnelles en question partitionnent les classes  $Z_{\text{anim}}^{\text{exp}}$  et  $Z_{\text{inan}}^{\text{exp}}$  respectivement en

8. C'est-à-dire l'ensemble des traits portés par une forme donnée, autrement dit une étiquette morphosyntaxique complète.

$z_{\text{anim,sg}}^{\text{exp}}$  et  $z_{\text{inan,sg}}^{\text{exp}}$ , correspondant aux réalisations des sous-paradigmes du singulier, d’une part et  $z_{\text{anim,pl}}^{\text{exp}}$  et  $z_{\text{inan,pl}}^{\text{exp}}$ , correspondant aux réalisations des sous-paradigmes du pluriel, de l’autre. Les zones flexionnelles associées aux noms d’animaux slovaques sont quant à elles les zones  $z_{\text{anim,sg}}^{\text{exp}}$  et  $z_{\text{inan,pl}}^{\text{exp}}$ , chacune de ces zones faisant par défaut partie d’une classe flexionnelle différente. L’appartenance à deux classes flexionnelles distinctes des zones flexionnelles associées à un lexème donné correspond à ce que nous appelons par définition HÉTÉROCLISE. Ce rapport entre classes flexionnelles et zones flexionnelles pour cet exemple du slovaque est schématiquement résumé par le tableau 4.3.

	$Z_{\text{anim}}^{\text{exp}}$ : MASC. ANIMÉ CHLAP « <i>homme, type</i> »		$Z_{\text{inan}}^{\text{exp}}$ : MASC. INANIMÉ DUB « <i>chêne</i> »		MASC. HÉTÉROCLITE OROL « <i>aigle</i> »	
	$z_{\text{anim,sg}}^{\text{exp}}$ : SG	$z_{\text{anim,pl}}^{\text{exp}}$ : PL	$z_{\text{inan,sg}}^{\text{exp}}$ : SG	$z_{\text{inan,pl}}^{\text{exp}}$ : PL	$z_{\text{anim,sg}}^{\text{exp}}$ : SG	$z_{\text{inan,pl}}^{\text{exp}}$ : PL
NOM	<i>chlap</i>	<i>chlap-i</i>	<i>dub</i>	<i>dub-y</i>	<i>orol</i>	<i>orl-y</i>
GEN	<i>chlap-a</i>	<i>chlap-ov</i>	<i>dub-a</i>	<i>dub-ov</i>	<i>orl-a</i>	<i>orl-ov</i>
DAT	<i>chlap-ovi</i>	<i>chlap-om</i>	<i>dub-u</i>	<i>dub-om</i>	<i>orl-ovi</i>	<i>orl-om</i>
ACC	<i>chlap-a</i>	<i>chlap-ov</i>	<i>dub</i>	<i>dub-y</i>	<i>orl-a</i>	<i>orl-y</i>
LOC	<i>chlap-ovi</i>	<i>chlap-och</i>	<i>dub-e</i>	<i>dub-och</i>	<i>orl-ovi</i>	<i>orl-och</i>
INS	<i>chlap-om</i>	<i>chlap-mi</i>	<i>dub-om</i>	<i>dub-mi</i>	<i>orl-om</i>	<i>orl-ami</i>

TABLE 4.3 : Hétéroclise dans les paradigmes des noms d’animaux slovaques

### 4.1.3 Classes de radicaux

Au delà de la notion de zone de radical introduite à la section précédente, PARSLI comporte également une formalisation de la notion de CLASSE DE RADICAUX.

Cette notion permet notamment de rendre compte des cas d’allomorphie radicale régulière, comme celle qui s’observe dans la flexion des verbes persan présentée au tableau 4.4<sup>9</sup>. Nous pouvons alors l’opposer aux cas de supplétion de radical où l’alternance radicale au sein d’un paradigme donné ne présente pas de régularité interne, comme dans le cas du verbe ALLER en français, illustré par un extrait de son paradigme au tableau 4.5.

Toutefois, cette distinction n’est pas faite dans toutes les approches théoriques récentes, et notamment dans celles concernant la flexion verbal du français. Bonami & Boyé (2003) récusent explicitement la distinction théorique entre les cas d’allomorphie radicale régulière, observables notamment pour les verbes du deuxième groupe en français comme FINIR, et les cas de supplétion de radicale du type de ce qui s’observe pour pour les verbes comme ALLER. Ils rendent ainsi compte des deux phénomènes par le même mécanisme.

9. Ce tableau rappelle le tableau 3.7 au chapitre 3.

LEXÈME	TRAD.	RAD1	RAD2
ÂRÂSTAN	« orner »	ârâst	ârâ
ÂMUXTAN	« apprendre »	âmuxt	âmuz
RAQSIDAN	« danser »	raqsid	raqs

TABLE 4.4 : Allomorphie radicale régulière (exemple des radicaux du présent (RAD1) et du passé (RAD2) en persan)

ALLER		
INF	<i>all-er</i>	al-e
1PL.PRS.IND	<i>nous all-ons</i>	nuz-al-õ
1SG.PRS.IND	<i>je v-ais</i>	ʒ(ə)-v-ε
2SG.PRS.IND	<i>tu v-as</i>	ty-v-a
1SG.PRS.SBJV	<i>(que)j'aill-e</i>	ʒ-aj
2SG.PRS.IND	<i>(que)tu aill-es</i>	ty-aj
1SG.FUT.IND	<i>j'irai</i>	ʒ-i-r-ε
2SG.FUT.IND	<i>tu iras</i>	ty-i-r-a

TABLE 4.5 : Supplétion de radicaux dans le paradigme du verbe ALLER

Si l'on utilise l'approche à douze radicaux de Bonami & Boyé (2003), les verbes du second groupe comme FINIR comporteront dans leur entrée lexicale un second radical en *-iss-* défini explicitement à côté du radical de base en *-i-*<sup>10</sup>. Nous donnons ci-dessous un exemple pour *finir*.

- 4.1 Analyse avec radicaux multiples (espaces thématiques (Bonami & Boyé, 2003)) :  
 exemple FINIR  
 Thème 1 : fini-  
 Thème 2 : finis-

Le verbe ALLER sera quant à lui défini par six thèmes différents<sup>11</sup> dans l'approche de Bonami & Boyé.

- 4.2 Analyse avec radicaux multiples (espaces thématiques (Bonami & Boyé, 2003)) :  
 exemple ALLER

10. FINIR ne distingue ainsi pas 12 radicaux, mais seulement deux dans l'espace des 12.

11. Bonami & Boyé (2003) distinguent deux thèmes différents pour *vont* et *va*.

- Thème 1 : all-
- Thème 2 : v̄-
- Thème 3 : va-
- Thème 7 : aj-
- Thème 10 : i-

Dans  $\mathcal{PARSL}$ , nous pouvons traiter les cas d'allomorphie régulière non pas par des radicaux lexicalement spécifiés, mais par l'emploi d'une zone de radical particulière par radical, définie dans la grammaire. De façon analogue à la définition  $\mathcal{PARSL}$  des classes flexionnelles au niveau des exposants, ces zones de radicaux, si elles sont fréquemment employées ensemble définiront ainsi une CLASSE DE RADICAUX spécifique.

Pour les verbes du deuxième groupe en français, une modélisation possible en  $\mathcal{PARSL}$  consistera ainsi à associer les lexèmes du deuxième groupe à deux zones de radicaux appartenant à une classe de radicaux unique. Ces zones de radicaux sont la zone  $z_{v2,1}^s$  pour les cases pour lesquelles les verbes du deuxième groupe utilisent un radical secondaire en *-iss-*, et la zone  $z_{v2,2}^s$  pour les autres cases utilisant le radical en *-i-*. Chacune de ces zones peut alors être associée à la classe flexionnelle (unique)  $Z_{v2}^{exp}$  des verbes du deuxième groupe. Le schème des verbes du second groupe est alors défini au moyen de deux sous-schémas (simplifiés), à savoir la zone de radicaux  $z_{v2,1}^s$  associée à la classe flexionnelle  $Z_{v2}^{exp}$  et la zone de radicaux  $z_{v2,2}^s$  combinée à  $Z_{v2}^{exp}$ . Les formes issues de chacun de ces sous-schémas sont à chaque fois les formes correspondant aux cases définies par les PTM compatibles de chacun des éléments d'un sous-schème donné : en d'autres termes, chaque sous-schème définit les réalisations des formes exprimant l'intersection des ensembles de PTM pour lesquels chacune de ses zones est définie. Cela évite notamment toute redondance, ou plus problématique encore, les cas de surgénération, dans la réalisation des formes d'un lexème donné.

Notre modélisation des verbes du deuxième groupe du français est illustrée pour le présent de l'indicatif aux figures 4.3 et 4.4.

Ce concept de classes de radicaux permet de distinguer les cas d'ALLOMORPHIE DE RADICAUX RÉGULIÈRE des cas extrêmes de SUPPLÉTION DE RADICAL. Dans un cas d'allomorphie régulière, le lexème est associé à une des zones de radicaux disponibles et la formation des radicaux se fait selon les règles de formation de radicaux spécifiés par la classe de radicaux. Les cas de supplétion de radical sont les cas où l'un des radicaux d'un lexème donné n'est pas formé régulièrement selon les règles de formation de radicaux disponibles, mais où son radical est listé spécifiquement parmi les informations morphologiques

Cl. de radicaux $Z_{v2}^s$			Cl. flexionnelle $Z_{v2}^{exp}$		
1SG		■ $z_{v2,1}^s$	1SG	-s	■ $Z_2^{exp}$
2SG	-i		2SG	-s	
3SG			3SG	-t	
1PL		■ $z_{v2,2}^s$	1PL	-ons	
2PL	-iss		2PL	-ez	
3PL			3PL	-ent	

FIGURE 4.3 : Classe et zones de radicaux (à gauche) et classe flexionnelle (à droite) du présent de l'indicatif des verbes du deuxième groupe en français

	$fini + Z_2^{exp}$	$finiss + Z_2^{exp}$
1SG	<i>finis</i>	
2SG	<i>finis</i>	
3SG	<i>finit</i>	
1PL		<i>finissons</i>
2PL		<i>finissez</i>
3PL		<i>finissent</i>

FIGURE 4.4 : Sous-schémas (simplifiés) pour le verbe *finir*

pertinentes pour le lexème en question, cf. l'exemple des radicaux *v-*, *aill-* et *ir-* du verbe ALLER en français qui coexistent avec le radical régulier *al-*. À la section 4.2, nous revenons plus précisément sur les implications et représentations formelles de cette distinction.

La figure 4.5 montre l'entrée lexicale partielle du verbe ALLER dans  $\mathcal{PARSLI}$ . Elle comporte une séquence phonologique initiale servant de point de départ à l'application des règles réalisationnelles noté I-PHON (pour inflectional PHONOlogical input en anglais), une catégorie flexionnelle I-CAT (pour inflectional CATegory) de type *verbe*, un inventaire de PTM MSF (pour MORPHO-syntactic FEatures), une liste S-STEM (pour suppletive STEM) de trois radicaux supplétifs assortis de leur identifiant de zone de radical et un schème flexionnel I-PAT (pour inflectional PATtern) que nous ne détaillons pas à ce stade. En l'occurrence, les traits exprimables par ALLER sont les même que ceux exprimés par sa catégorie. Nous avons rendu ce fait par la notation *standard*<sup>12</sup>.

La figure 4.6 indique l'entrée lexicale correspondant au verbe FINIR. À la différence de l'entrée pour ALLER elle ne comporte pas de liste de radicaux supplétifs. En revanche, son schème (I-PAT) flexionnel indique l'utilisation de deux zones de radicaux distinctes pour la formation du paradigme de FINIR. Ces zones sont indiquées au sein de couples réalisationnels constitués pour chacun d'entre eux d'une classe réalisationnelle (comme la

12. Plutôt que de lister l'intégralité des PTM canoniquement exprimés par les verbes en français.

ALLER	
I-PHON	aller
I-CAT	<i>verbe</i>
MSF	{ <i>standard</i> }
S-STEM	$z_1^s$ : <i>v-</i> $z_3^s$ : <i>aill-</i> $z_4^s$ : <i>i-</i>
S-FORM	( <i>vide</i> )
I-PAT	...

FIGURE 4.5 : Entrée lexicale pour le verbe ALLER

classe flexionnelle  $Z_{v2}^{exp}$ ) ou une zone réalisationnelle (comme la zone de radical  $z_{v2,1}^s$ ) et d'une fonction de transfert id sur laquelle nous reviendrons plus bas.

FINIR	
I-PHON	finir
I-CAT	<i>verbe</i>
MSF	{ <i>standard</i> }
S-STEM	( <i>vide</i> )
S-FORM	( <i>vide</i> )
I-PAT	$(z_{v2,1}^s, id), (Z_{v2}^{exp}, id)$ $(z_{v2,2}^s, id), (Z_{v2}^{exp}, id)$

FIGURE 4.6 : Entrée lexicale pour le verbe FINIR

Les vrais cas de supplétion radicale se distinguent ainsi des cas d'allomorphie radicale régulière par une répartition différente de l'information entre la grammaire morphologique et le lexique morphologique. Dans le cas de l'allomorphie régulière, l'alternance radicale est prise en charge par des zones de radicaux appartenant à la grammaire et associées à un lexème donné par un schème flexionnel (I-PAT) particulier. Dans le cas de la supplétion de radicaux, les radicaux supplétifs sont explicitement listés dans l'entrée lexicale concernée (sous S-STEM).

Notons enfin que dans un cas de SUPPLÉTION DE FORME dans le paradigme d'un

lexème donné, ce n'est pas seulement le radical mais l'intégralité de sa forme qui est listée spécifiquement parmi les informations morphologiques pertinentes, cf. l'exemple du 1PL.PRS.IND du verbe ÊTRE « *sommes* » que nous avons présenté au chapitre 3.

La figure 4.7 montre ainsi l'entrée lexicale partielle du verbe ÊTRE dans  $\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}\mathbb{I}$ . Elle comporte sa séquence phonologique initiale I-PHON<sup>13</sup>, une catégorie flexionnelle I-CAT de type *verbe*, un inventaire de PTM standard, une liste S-FORM<sup>14</sup> de trois formes supplétives assorties de leur PTM exprimé et un schème flexionnel I-PAT que nous ne détaillons pas à ce stade.

ÊTRE		
I-PHON	être	
I-CAT	<i>verbe</i>	
MSF	{ <i>standard</i> }	
S-STEM	( <i>vide</i> ) }	
S-FORM	1SG.PRS.IND <i>sommes</i>	
	1PL.PRS.IND <i>suis</i>	
	3PL.PRS.IND <i>sont</i>	
I-PAT	...	

FIGURE 4.7 : Entrée lexicale pour le verbe ÊTRE

#### 4.1.4 Le niveau thématique

En plus de ces deux dimensions associées respectivement aux radicaux et aux exposants,  $\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}\mathbb{I}$  permet de définir un NIVEAU THÉMATIQUE, lui aussi découpé en ZONES THÉMATIQUES. Le niveau thématique participe à la formation du radical au sens de Matthews (1974), mais il se distingue du niveau radical strict par le fait que sa réalisation est optionnelle<sup>15</sup>. Le niveau radical est quant à lui obligatoire au sens où tout mot-forme

13. Par simplicité nous avons dans les exemples de ce chapitre pris la forme de citation (l'infinitif) comme valeur de I-PHON. Ce choix est évidemment contestable pour le français. Il a été retenu dans l'intérêt de la cohérence avec les analyses du lexique électronique à grande échelle Lefff Sagot (2010) sur lequel s'appuient nos analyses du français présentées au chapitre 6. Le fait de poser la forme de citation habituelle comme contenu de I-PHON n'est cependant pas obligatoire. Il serait envisageable d'utiliser un radical, voire une racine partiellement reconstruite. Le choix de la valeur de I-PHON se répercutera cependant nécessairement sur l'inventaire des règles de réalisation à spécifier pour la construction des formes.

14. Pour suppletive STEM en anglais.

15. Notons que notre emploi du terme « *thème* » ici ne se limite pas à la notion de thème utilisée pour caractériser la la sélection des thèmes (*stem*) dans la flexion verbale dans les langues romanes au sens de

comporte au moins une réalisation de niveau radical<sup>16</sup>.

Nous avons introduit le niveau thématique au chapitre 1 pour rendre compte des verbes latins qui affichent une appartenance partielle à la classe des verbes à thème en *i* d'une part et aux verbes athématiques de l'autre. Nous rappelons la répartition de l'appartenance thématique aux tableaux 4.6 (verbes à thème en *i*), 4.7 (verbes athématiques) et 4.8 (verbes à thème mixte).

Les verbes latins sont ainsi chacun associés à un thème. Dans les exemples indiqués dans les tableaux 4.6 et 4.7, les verbes appartiennent l'un au thème en *i*, que nous notons ici  $Z_i^{th}$ , et l'autre à la classe athématique, notée ici  $Z_{none}^{th}$ . Les verbes comme ACCIPIO « *accepter* » et ACCIPIOR « *être accepté* » quant à eux sont associés à deux zones thématiques chacun, appartenant chacune à un thème différent : une zone  $z_{i,1}^{th}$  appartenant au thème  $Z_i^{th}$  des verbes à thème en *i* et une  $z_{none,2}^{th}$  appartenant au thème  $Z_{none}^{th}$  des verbes athématiques. Cette façon de représenter l'appartenance thématique des verbes mixtes à deux thèmes par l'intermédiaire de zones thématiques nous permet de définir dans  $\mathcal{PARSL}$  un cas d'HÉTÉROCLISE THÉMATIQUE analogue à la notion d'hétéroclise au niveau exposant.  $\mathcal{PARSL}$  permet ainsi de définir les phénomènes non canoniques niveau par niveau et de représenter explicitement à quel niveau de la réalisation de la forme intervient un phénomène non canonique particulier. Cette représentation permet aussi de rendre explicite ce que montrent des données de paradigmes comme celui d'ACCIPIO, à savoir que le phénomène de l'hétéroclise ne se limite pas à l'exponence.

## 4.1.5 Une représentation multi-niveaux

### 4.1.5.1 Un « mille-feuille » réalisationnel

D'une manière générale, le volet réalisationnel de  $\mathcal{PARSL}$  prévoit une représentation de la réalisation flexionnelle à plusieurs niveaux. Nous avons introduit les trois types de niveaux : le NIVEAU RADICAL, le NIVEAU THÉMATIQUE et le NIVEAU FLEXIONNEL ou NIVEAU D'EXPONENCE. Chacun comporte respectivement une association des lexèmes avec un ensemble de zones de radicaux, de zones thématiques et de zones flexionnelles. Chaque zone peut à son tour être par défaut associée à une classe de radicaux, un thème ou une classe flexionnelle.

Roché (2010) ou encore Bonami & Boyé (2003). Comme nous l'avons indiqué au chapitre 1, la notion que nous introduisons ici désigne un niveau d'analyse intermédiaire entre les radicaux et les exposants. Il est susceptible de pouvoir être employé dans la description de toute langue et ne se limite par conséquent pas à la qualification d'un sous-ensemble de systèmes flexionnels, cf. notre description du maltais au chapitre 5.

16. Pour les mots invariables, cette réalisation est identique à la forme de citation.



		AUDIO						
		ACT		PASS				
		IND	PL	IND	PL	IND	PL	
		SG	PL	SG	PL	SG	PL	
IPFV	PRS	1	<i>audi-ō</i>	<i>audi-am</i>	<i>audi-ōr</i>	<i>audi-mur</i>	<i>audi-ar</i>	<i>audi-āmur</i>
		2	<i>audi-s</i>	<i>audi-ās</i>	<i>audi-ris (-re)</i>	<i>audi-minī</i>	<i>audi-āris (-āre)</i>	<i>audi-āminī</i>
		3	<i>audi-t</i>	<i>audi-at</i>	<i>audi-tur</i>	<i>audi-ntur</i>	<i>audi-antur</i>	<i>audi-antur</i>
PST	1	<i>audi-ēbam</i>	<i>audi-ēbamus</i>	<i>audi-ēbar</i>	<i>audi-ēbāmur</i>	<i>audi-ēbamur</i>	<i>audi-ēbāmur</i>	
		2	<i>audi-ēbas</i>	<i>audi-rēs</i>	<i>audi-ēbāris (-ēbāre)</i>	<i>audi-ēbāminī</i>	<i>audi-rēris (-rēre)</i>	<i>audi-rēminī</i>
		3	<i>audi-ēbat</i>	<i>audi-rēt</i>	<i>audi-ēbātur</i>	<i>audi-ēbantur</i>	<i>audi-rētur</i>	<i>audi-rētur</i>
FUT	1	<i>audi-am</i>	<i>audi-erimus</i>	<i>audi-ar</i>	<i>audi-erimur</i>	<i>audi-erimur</i>	<i>audi-erimur</i>	
		2	<i>audi-ēs</i>	<i>audi-eritis</i>	<i>audi-eris (-ere)</i>	<i>audi-erimī</i>	<i>audi-eritis (-erere)</i>	<i>audi-erimī</i>
		3	<i>audi-et</i>	<i>audi-erint</i>	<i>audi-entur</i>	<i>audi-erimī</i>	<i>audi-erintur</i>	<i>audi-erintur</i>
PFV	PRS	1	<i>audi-o-ī</i>	<i>audi-o-erim</i>	<i>audi-us sum</i>	<i>audi-t-ī sumus</i>	<i>audi-us sim</i>	<i>audi-t-ī simus</i>
		2	<i>audi-o-istī</i>	<i>audi-o-eritis</i>	<i>audi-us es</i>	<i>audi-t-ī estis</i>	<i>audi-us sis</i>	<i>audi-t-ī stis</i>
		3	<i>audi-o-it</i>	<i>audi-o-erit</i>	<i>audi-us est</i>	<i>audi-t-ī sunt</i>	<i>audi-us sit</i>	<i>audi-t-ī stit</i>
PST	1	<i>audi-o-eram</i>	<i>audi-o-erāmus</i>	<i>audi-o-eram</i>	<i>audi-t-ī eramus</i>	<i>audi-us essem</i>	<i>audi-t-ī essemus</i>	
		2	<i>audi-o-erās</i>	<i>audi-o-erātis</i>	<i>audi-us erās</i>	<i>audi-t-ī erātis</i>	<i>audi-us essēs</i>	<i>audi-t-ī essētis</i>
		3	<i>audi-o-erat</i>	<i>audi-o-erant</i>	<i>audi-us erat</i>	<i>audi-t-ī erant</i>	<i>audi-us esset</i>	<i>audi-t-ī essent</i>
FUT	1	<i>audi-o-erō</i>	<i>audi-o-erimus</i>	<i>audi-us erō</i>	<i>audi-t-ī erimus</i>	<i>audi-us erim</i>	<i>audi-t-ī erimus</i>	
		2	<i>audi-o-eris</i>	<i>audi-o-eritis</i>	<i>audi-us eris</i>	<i>audi-t-ī eritis</i>	<i>audi-us eritis</i>	<i>audi-t-ī eritis</i>
		3	<i>audi-o-erit</i>	<i>audi-o-erint</i>	<i>audi-us erit</i>	<i>audi-t-ī erint</i>	<i>audi-us erint</i>	<i>audi-t-ī erint</i>
IMP	PRS	2	<i>audi-te</i>	<i>audi-te</i>	<i>audi-te</i>	<i>audi-minī</i>	<i>audi-te</i>	<i>audi-minī</i>
		2	<i>audi-to</i>	<i>audi-tōte</i>	<i>audi-tor</i>	<i>audi-ntor</i>	<i>audi-tor</i>	<i>audi-ntor</i>
		3	<i>audi-to</i>	<i>audi-tōte</i>	<i>audi-tor</i>	<i>audi-ntor</i>	<i>audi-tor</i>	<i>audi-ntor</i>
INF	PRS	<i>audi-re</i>		<i>audi-re</i>		<i>audi-ri</i>		
		<i>audi-to</i>		<i>audi-tōte</i>		<i>audi-ntor</i>		
		<i>audi-to</i>		<i>audi-tōte</i>		<i>audi-ntor</i>		
PTCP	PST	<i>audi-t-ūrus/-ūra/-ūrum</i>		<i>audi-t-ūrus/-ūra/-ūrum esse</i>		<i>audi-t-ūrus/-ūra/-ūrum esse</i>		
		<i>audi-t-ūrus/-ūra/-ūrum esse</i>		<i>audi-t-ūrus/-ūra/-ūrum esse</i>		<i>audi-t-ūrus/-ūra/-ūrum esse</i>		
		<i>audi-t-ūrus/-ūra/-ūrum esse</i>		<i>audi-t-ūrus/-ūra/-ūrum esse</i>		<i>audi-t-ūrus/-ūra/-ūrum esse</i>		
GRV	GRD	<i>audi-t-ūrus/-ūra/-ūrum</i>		<i>audi-t-ūrus/-ūra/-ūrum</i>		<i>audi-t-ūrus/-ūra/-ūrum</i>		
		<i>audi-t-ūrus/-ūra/-ūrum</i>		<i>audi-t-ūrus/-ūra/-ūrum</i>		<i>audi-t-ūrus/-ūra/-ūrum</i>		
		<i>audi-t-ūrus/-ūra/-ūrum</i>		<i>audi-t-ūrus/-ūra/-ūrum</i>		<i>audi-t-ūrus/-ūra/-ūrum</i>		
SUP	FUT	<i>audi-ndus/-nda/-ndum</i>		<i>audi-ndus/-nda/-ndum</i>		<i>audi-ndus/-nda/-ndum</i>		
		<i>audi-ndus/-nda/-ndum</i>		<i>audi-ndus/-nda/-ndum</i>		<i>audi-ndus/-nda/-ndum</i>		
		<i>audi-ndus/-nda/-ndum</i>		<i>audi-ndus/-nda/-ndum</i>		<i>audi-ndus/-nda/-ndum</i>		

TABLE 4.6 : Voyelle thématique *i* latins — exemples AUDIO « entendre » et AUDIOR « être entendu »

		DUCO							
		ACT			PASS				
IPFV	PRS	IND		SBJV		IND		SBJV	
		SG	PL	SG	PL	SG	PL	SG	PL
PFV	1	<i>dūc-ō</i>	<i>dūci-mus</i>	<i>dūc-am</i>	<i>dūc-āmus</i>	<i>dūc-or</i>	<i>dūci-mur</i>	<i>dūc-ar</i>	<i>dūc-āmur</i>
	2	<i>dūc-t-s</i>	<i>dūci-tis</i>	<i>dūc-ās</i>	<i>dūc-ātis</i>	<i>dūc-ris (-re)</i>	<i>dūci-mini</i>	<i>dūc-aris (-āre)</i>	<i>dūc-āmini</i>
	3	<i>dūc-t</i>	<i>dūcu-nt</i>	<i>dūc-at</i>	<i>dūc-ant</i>	<i>dūci-tur</i>	<i>dūc-untur</i>	<i>dūc-ātur</i>	<i>dūc-antur</i>
	1	<i>dūc-ēbam</i>	<i>dūc-ēbamus</i>	<i>dūc-rem</i>	<i>dūc-rēmus</i>	<i>dūc-ēbar</i>	<i>dūc-ēbāmur</i>	<i>dūc-er</i>	<i>dūc-rēmur</i>
	2	<i>dūc-ēbās</i>	<i>dūc-ēbātis</i>	<i>dūc-rēs</i>	<i>dūc-rētis</i>	<i>dūc-ēbāris (-ēbare)</i>	<i>dūc-ēbāmini</i>	<i>dūc-rēris (-rēre)</i>	<i>dūc-rēmini</i>
	3	<i>dūc-ēbat</i>	<i>dūc-ēbant</i>	<i>dūc-ret</i>	<i>dūc-rent</i>	<i>dūc-ēbātur</i>	<i>dūc-ēbantur</i>	<i>dūc-rētur</i>	<i>dūc-rentur</i>
	1	<i>dūc-am</i>	<i>dūc-ēmus</i>	<i>dūc-ret</i>	<i>dūc-rent</i>	<i>dūc-ar</i>	<i>dūc-ēmur</i>		
	2	<i>dūc-ēs</i>	<i>dūc-ētis</i>			<i>dūc-eris (-ēre)</i>	<i>dūc-ēmini</i>		
	3	<i>dūc-et</i>	<i>dūc-ent</i>			<i>dūc-ētur</i>	<i>dūc-entur</i>		
IMP	1	<i>dūc-ī</i>	<i>dūc-imus</i>	<i>dūc-erim</i>	<i>dūc-erimus</i>	<i>dūc-us sum</i>	<i>dūc-ī sumus</i>	<i>dūc-us sim</i>	<i>dūc-ī simus</i>
	2	<i>dūc-istī</i>	<i>dūc-istis</i>	<i>dūc-eris</i>	<i>dūc-eritis</i>	<i>dūc-us es</i>	<i>dūc-ī estis</i>	<i>dūc-us sis</i>	<i>dūc-ī stis</i>
	3	<i>dūc-it</i>	<i>dūc-erunt (-ere)</i>	<i>dūc-erit</i>	<i>dūc-erint</i>	<i>dūc-us est</i>	<i>dūc-ī sunt</i>	<i>dūc-us sit</i>	<i>dūc-ī sint</i>
	1	<i>dūc-eram</i>	<i>dūc-erāmus</i>	<i>dūc-isse</i>	<i>dūc-issēmus</i>	<i>dūc-us eram</i>	<i>dūc-ī erāmus</i>	<i>dūc-us essem</i>	<i>dūc-ī essemus</i>
	2	<i>dūc-erās</i>	<i>dūc-erātis</i>	<i>dūc-issēs</i>	<i>dūc-issētis</i>	<i>dūc-us erās</i>	<i>dūc-ī erātis</i>	<i>dūc-us essēs</i>	<i>dūc-ī essētis</i>
	3	<i>dūc-erat</i>	<i>dūc-erant</i>	<i>dūc-isset</i>	<i>dūc-issent</i>	<i>dūc-us erat</i>	<i>dūc-ī erant</i>	<i>dūc-us esset</i>	<i>dūc-ī essent</i>
	1	<i>dūc-erō</i>	<i>dūc-erimus</i>			<i>dūc-us erō</i>	<i>dūc-ī erimus</i>		
	2	<i>dūc-eris</i>	<i>dūc-eritis</i>			<i>dūc-us eris</i>	<i>dūc-ī eritis</i>		
	3	<i>dūc-erit</i>	<i>dūc-erint</i>			<i>dūc-us erit</i>	<i>dūc-ī erint</i>		
INF	PRS	<i>dūc</i>	<i>dūci-te</i>	<i>dūc-re</i>	<i>dūc-ī</i>	<i>dūc-re</i>	<i>dūci-mini</i>		
	FUT	<i>dūci-to</i>	<i>dūci-tōte</i>	<i>dūc-isse</i>	<i>dūc-ī</i>	<i>dūci-tor</i>	<i>dūc-ntor</i>	<i>dūc-us/-a/-um esse</i>	
		<i>dūci-to</i>	<i>dūci-tōte</i>	<i>dūc-ens, -entis</i>					
PTCP	PRS			<i>dūc-ens, -entis</i>					
	PST								
	FUT			<i>dūc-urus/-ūra/-ūrum esse</i>					
GRV	PRS								
	PST								
	FUT			<i>dūc-urus/-ūra/-ūrum</i>					
GRD				<i>dūc-endū/-endō/-endum</i>					
				<i>dūc-endus/-enda/-endum</i>					
SUP				<i>dūc-um/-ū</i>					

TABLE 4.7 : Verbes athématiques latins — exemples DUCO « conduire, mener » et DUCOR « être conduit »

				ACCIPIO						
				ACT		PASS				
				IND	SBJV	IND	PASS			
				SG	PL	SG	SBJV			
				IND	SBJV	IND	PL			
IPFV	PRS	1	accip-ō	accip-ōnus	accip-ānus	accip-ōr	accip-īmur	accip-ār	accip-īamur	
		2	accip-īs	accip-ītis	accip-ās	accip-ātis	accip-eris (ēre)	accip-īminī	accip-āris (āre)	accip-āminī
		3	accip-it	accip-unt	accip-āt	accip-ant	accip-itur	accip-untur	accip-antur	accip-antur
PST	1	1	accip-ebam	accip-ebānus	accip-erem	accip-erēmus	accip-ēbar	accip-ēbāmur	accip-erem	accip-erēmur
		2	accip-ebās	accip-ebātis	accip-erēs	accip-erētis	accip-ēbāris (āre)	accip-ēbāminī	accip-erēs (ere)	accip-erēmī
		3	accip-ebat	accip-ebant	accip-eret	accip-erent	accip-ēbatur	accip-ēbantur	accip-eretur	accip-erentur
FUT	1	1	accip-am	accip-ēmus		accip-āur	accip-ēmur			
		2	accip-ēs	accip-ētis		accip-eris (ēre)	accip-ēmīnī			
		3	accip-et	accip-ent		accip-ētur	accip-entur			
PFV	PRS	1	accēp-ī	accēp-īnus	accēp-erim	accēp-erimus	accēp-īsum	accēp-īsumus	accēp-īsimus	
		2	accēp-istī	accēp-istis	accēp-eris	accēp-eritis	accēp-īes	accēp-īes	accēp-īes	
		3	accēp-it	accēp-erunt (ēre)	accēp-erit	accēp-erint	accēp-īes	accēp-īes	accēp-īes	
PST	1	1	accēp-eram	accēp-erānus	accēp-erem	accēp-erēmus	accēp-īeram	accēp-īerāmus	accēp-īessem	
		2	accēp-erās	accēp-erātis	accēp-erēs	accēp-erētis	accēp-īerās	accēp-īerātis	accēp-īesēs	
		3	accēp-erat	accēp-erant	accēp-erit	accēp-erint	accēp-īerat	accēp-īerant	accēp-īesset	
FUT	1	1	accēp-erō	accēp-erimus		accēp-īerō	accēp-īerimus			
		2	accēp-eris	accēp-eritis		accēp-īeris	accēp-īeritis			
		3	accēp-erit	accēp-erint		accēp-īerit	accēp-īerint			
IMP	PRS	2	accipe	accip-ite		accipe-re	accip-imini			
		2	accip-to	accip-ito		accip-tor	accip-itor			
		3	accip-to	accip-ito		accip-tor	accip-itor			
INF	PRS			accipe-re			accip-ī			
		PST		accēp-isse						
		FUT		accēp-ūrus/-ūral/-ūrum esse						
PTCP	PRS			accip-ēns, -ēntis			accēp-ns/-al/-um			
		PST								
		FUT								
GRV GRD SUP	PST			accēp-ūrus/-ūral/-ūrum						
				accip-ēndī/-endō/-endum						
				accip-ēndus/-enda/-endum						
SUP	PST			accēp-ūrus/-ūral/-ūrum						
				accip-ēndus/-enda/-endum						
				accēp-ūrus/-ūral/-ūrum						
				athématique						

TABLE 4.8 : Verbes latins à thème mixte — exemples ACCIPIO « *accipere* » et ACCIPIOR « *être accipere* »

Les propriétés distinctives de ces trois types de niveaux sont les suivantes :

- Le niveau radical est obligatoire : toute réalisation de forme comporte au moins la réalisation d'un radical. Un cas extrême est celui des mots canoniquement invariables (comme les conjonctions de coordination ou les prépositions en français) qui réalisent uniquement un même radical à travers l'intégralité de leur paradigme — qui en l'occurrence consiste en une case unique. Les zones de radicaux sont sélectionnées en fonction des traits à réaliser par la forme. Elles contiennent chacune une fonction de formation de radical.

Dans la notation  $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$  habituelle, une zone de radicaux est généralement associée à la notation  $z^s$ , une classe de radicaux à la notation  $Z^s$ .

- Le niveau thématique est optionnel. Il comprend des zones thématiques appartenant à un thème donné. Un lexème donné peut être associé à une zone thématique donnée pour la réalisation d'une forme particulière, c'est-à-dire en fonction des traits morphosyntaxiques que réalise la forme en question.

Une zone thématique est par convention notée  $z^{th}$ , un thème (ou classe thématique) est noté  $Z^{th}$ .

- Le niveau d'exponence (ou flexionnel) est lui aussi optionnel. Lorsqu'il est présent il représente nécessairement la réalisation des traits morphosyntaxiques exprimés par la forme. Il comporte des zones flexionnelles qui peuvent être associées à des classes flexionnelles. Chaque zone flexionnelle contient un ensemble de fonctions ou règles de réalisation.

En  $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$ , on représente généralement par  $z^{exp}$  une zone flexionnelle et par  $Z^{exp}$  une classe flexionnelle.

Pour illustration, nous rappelons à la figure 4.8 le schéma représentant une analyse en trois niveaux de la flexion latine que nous avons donnée au chapitre 1. Les niveaux représentés sont respectivement le niveau radical (en rouge), le niveau thématique (en vert) et le niveau exposant (en bleu).

Le type d'information morphophonologique<sup>17</sup> peut néanmoins varier d'une langue à l'autre. Par ailleurs, les différents types de niveaux (radical, thématique, exponence) peuvent également être représentés par plusieurs niveaux individuels pour un type de niveau donné dans une langue donnée. Cette représentation s'inspire de la représentation de la flexion des verbes en chinantec de Tlatepuzco<sup>18</sup> par Palancar (à paraître) : chaque

17. Par exemple des informations segmentales par opposition à des informations suprasegmentales de type ton ou accentuation.

18. Langue otomangue parlée dans l'état d'Oaxaca, dans le sud du Mexique.

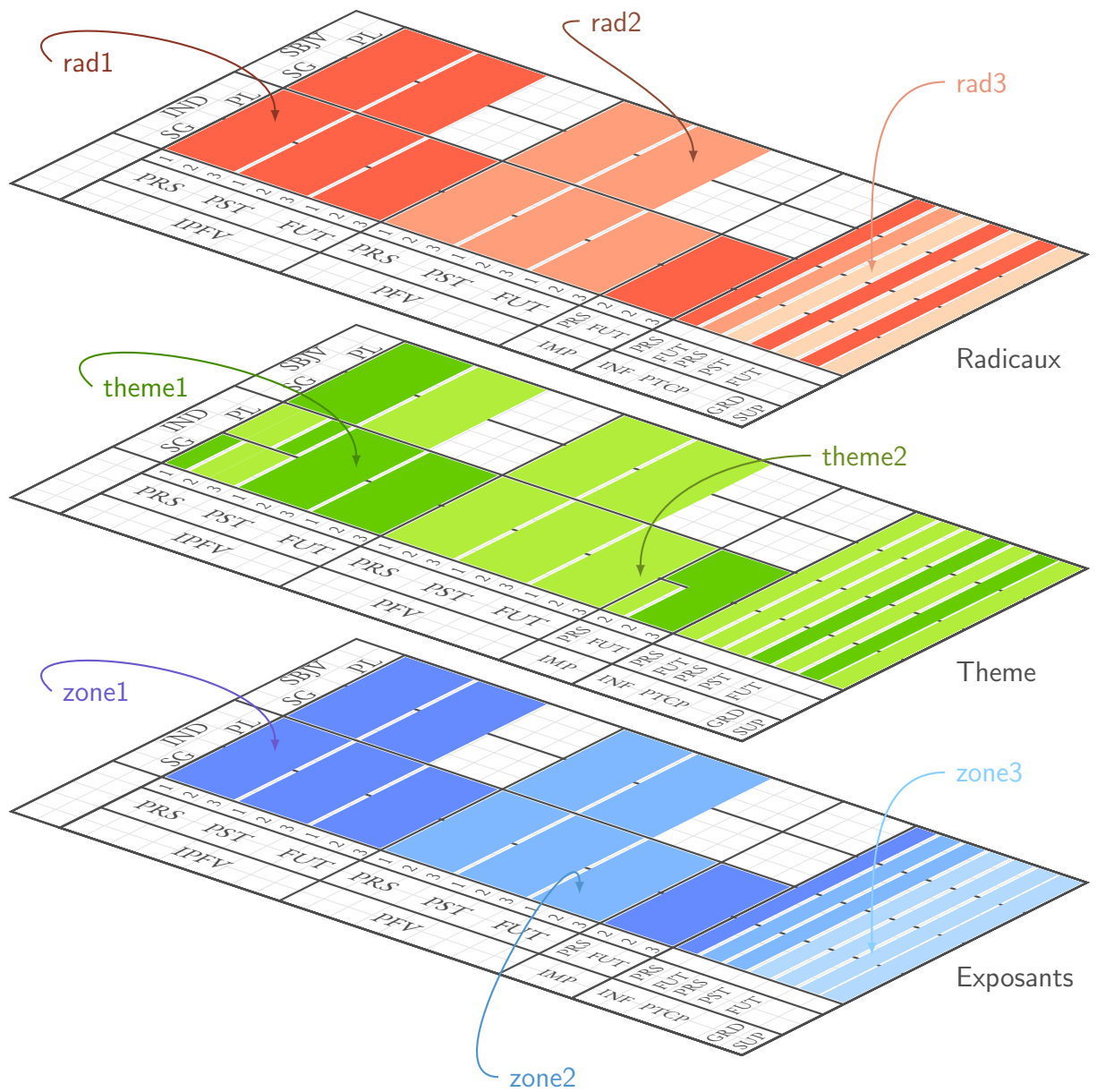


FIGURE 4.8 : Représentation des niveaux réalisationnels dans  $\mathcal{PARSLI}$  (exemple de l'analyse du latin donnée au chapitre 1)

forme  $y$  est définie par des propriétés segmentales (forme du radical et préfixes TAM) et des propriétés suprasegmentales (ton, type d'accent). Nous montrons des extraits de deux paradigmes du chinantec au tableau 4.9. Les données sont empruntées à Palancar (c.p.); elles se concentrent ici sur l'interaction entre tons et accents. Les tons  $y$  sont représentés par des chiffres en exposant : 1 pour un ton bas, 2 pour un ton moyen et 3 pour un ton haut, ainsi que des séquences de ces nombres pour les tons modulés<sup>19</sup>. Les accents sont de deux types : un accent dit « contrôlé » et un accent dit « ballistique ». Pour simplifier les définitions qui ont pu être données de l'accent ballistique, nous dirons ici qu'il correspond à une articulation plus marquée que l'accent contrôlé. L'accent ballistique est noté par un accent aigu typographique dans les exemples du tableau 4.9.

	JÚH <sup>2</sup> « courber »				TIU <sup>12</sup> « couper »			
PERSONNE	3	1SG	2	1PL	3	1SG	2	1PL
RAD.PRS	júb <sup>2</sup>	júb <sup>12</sup>	júb <sup>12</sup>	júb <sup>12</sup>	tiu <sup>12</sup>	tiu <sup>12</sup>	tiu <sup>2</sup>	tiu <sup>2</sup>
RAD.FUT	júb <sup>2</sup>	júb <sup>13</sup>	júb <sup>13</sup>	júb <sup>13</sup>	tiu <sup>1</sup>	tiu <sup>13</sup>	tiu <sup>3</sup>	tiu <sup>3</sup>
RAD.COMPL	júb <sup>2</sup>	júb <sup>1</sup>	júb <sup>1</sup>	júb <sup>13</sup>	tiu <sup>1</sup>	tiu <sup>1</sup>	tiu <sup>32</sup>	tiu <sup>3</sup>

TABLE 4.9 : Exponence multiple en chinantec de Tlatepuzco (données de Palancar (à paraître))

Les formes verbales comportent donc deux types d'exposants suprasegmentaux (ton et type d'accent). Chaque type d'exposant peut être pris en charge par un niveau d'exponence distinct dans le volet réalisationnel de  $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$ . Une telle représentation permet de tirer des généralisations séparées sur les classes d'exposants tonaux ou accentuels. Le volet réalisationnel de  $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$  peut ainsi être comparé à un mille-feuille comportant autant de niveaux que de procédés intervenant dans l'exponence ou la formation des radicaux. Les cas d'EXPONENCE MULTIPLE se traduisent ainsi par l'existence de plus d'un niveau d'exponence dans la réalisation d'un paradigme pour un système flexionnel donné. Par ailleurs, les mesures que nous présentons plus bas peuvent également être appliquées à chacun de ces niveaux.

Il est à noter que les niveaux n'ont pas toujours de raison théorique *a priori* d'être ordonnés entre eux. Par exemple, l'accent et le patron tonal semblent indépendants l'un de l'autre en chinantec. Toutefois, le formalisme d'implémentation que nous présenterons au chapitre suivant (chapitre 5) est conçu pour appliquer les règles de réalisation les unes après

19. 12 pour un ton modulé bas vers moyen, 13 pour un ton modulé de bas vers haut et 32 pour un ton modulé haut vers moyen.

les autres et nécessite donc de fait de lister les niveaux dans un certain ordre. En revanche, les trois types de niveaux (radical, thème et exponence) sont effectivement conçus pour être réalisés successivement : en fonction du paquet de traits morphosyntaxiques, le radical est en un premier temps formé au niveau radical ; de façon optionnelle, il peut être étendu au niveau thématique ; les exposants sont enfin réalisés sur le radical (étendu) par les règles de réalisation stipulées au niveau d'exponence.

#### 4.1.5.2 Les objets construits

Les différents niveaux de  $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$  ne construisent pas tous le même type d'objet. Les niveaux radical et thématiques combinés permettent de construire ce qui correspond au radical chez Matthews (1974). Les niveaux d'exponence combinés réalisent une forme complète à partir d'un radical donné. Cette représentation permet d'identifier la SUPPLÉTION DE RADICAUX aux cas où les niveaux de radicaux sont court-circuités dans la réalisation flexionnelle et lexicalement remplacés par un radical complet. La SUPPLÉTION DE FORME au sens de Boyé (2006), quant à elle, correspond aux cas analogues où tous les niveaux (radical, thématiques et d'exponence) sont court-circuités : le lexique spécifie alors une forme complète à la place de la réalisation régulière.

Dans le formalisme associé à  $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$ , nous notons  $s$  un radical entièrement formé et  $f$  un mot-forme entièrement réalisé. Les SYNCRÉTISMES peuvent naturellement se mesurer à l'égalité formelle de deux objets construits, qu'ils soient des radicaux ou des formes complètes.

#### 4.1.5.3 Blocs de règles et niveaux réalisationnels

Les nouveautés en termes d'encodage du processus réalisationnel apportées par  $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$  sont ainsi les suivantes :

- une généralisation des partitions du paradigme telles qu'elles ont été proposées par Pirrelli & Battista (2000) et Bonami & Boyé (2003) pour les radicaux à l'ensemble du mécanisme flexionnel, c'est-à-dire au niveau d'exponence par l'introduction du concept de ZONE FLEXIONNELLE ou de ZONE D'EXPONENCE, et au niveau thématique par la notion de ZONE THÉMATIQUE ;
- les classes flexionnelles ne sont pas définies *a priori*, mais dérivées des associations de zones flexionnelles observables de façon récurrente dans les ensembles de zones flexionnelles construisant les paradigmes des lexèmes de la langue,

- une représentation de la flexion comme une combinaison de règles identifiables à différents NIVEAUX RÉALISATIONNELS ;
- chacun de ces niveaux est structuré en zones appelées ZONES RÉALISATIONNELLES et pouvant être des zones de radicaux, des zones thématiques ou des zones flexionnelles/d'exponence ;
- les ensembles de PTM définissant à chaque niveau les différentes zones réalisationnelles sont appelés ESPACES PARTITIONNANTS<sup>20</sup>
- la notion de classe est définie pour chaque niveau réalisationnel comme une association par défaut de zones réalisationnelles ; nous appelons une telle classe CLASSE RÉALISATIONNELLE ;
- la combinaison de ces zones de niveaux différents pour la construction d'un paradigme est précisée par des SCHÈMES et des SOUS-SCHÈMES associés aux entrées lexicales flexionnelles individuelles et indiquant leur mode de flexion.

Notons que dans un niveau donné les règles de réalisation peuvent être structurées en blocs de règles tels que définis par Stump (2001)<sup>21</sup>. Ces blocs permettent de réaliser successivement des parties d'exponence sur les mots-formes. Mais les fonctionnements des blocs d'une part et des niveaux de l'autre présentent une différence fondamentale en PARSLI. La différence entre des blocs de règles et des niveaux réalisationnels est ainsi la suivante :

- Poser deux niveaux réalisationnels différents permet de définir des classes réalisationnelles dans chacun de ses niveaux ; les blocs sont internes aux zones et ainsi sans lien direct avec la notion de classe réalisationnelle ;
- Les classes de niveaux différents comportent des zones réalisationnelles définies par des espaces partitionnants différents ; les blocs différents à l'intérieur d'une zone sont des regroupements de règles de réalisation analogues au système de blocs en PFM ; les blocs à l'intérieur d'un niveau donné regroupent les règles de réalisation mutuellement exclusives à l'intérieur d'un même niveau ;
- Si deux ensembles d'exposants sont strictement co-variants mais nécessitent d'être réalisés séparément<sup>22</sup>, il ne relèvent pas de niveaux réalisationnels différents (qui

---

20. Une zone réalisationnelle correspond ainsi à l'intersection entre un espace partitionnant et un niveau réalisationnel.

21. Voir le chapitre 5 pour la différence d'implémentation des niveaux et des blocs dans le formalisme d'implémentation Alexina<sub>PARSLI</sub> dédié à l'implémentation de descriptions PARSLI.

22. C'est-à-dire qu'ils seraient nécessairement traités dans des blocs différents en PFM.



- pourraient donner lieu à des découpages en zones selon des espaces partitionnants différents) mais bien de blocs de règles distincts à l'intérieur d'un même niveau ;
- La fonction première des blocs est de permettre de rendre compte de classes de positions ; la fonction première des niveaux est de permettre des découpages en zones différents et de rendre compte des généralisations de type classe réalisationnelle lorsqu'elles sont indépendantes niveau par niveau ;
  - Le concept de RÈGLE PORTE-MANTEAU concerne exclusivement les règles de réalisation au sein d'une zone réalisationnelle donnée à un niveau réalisationnel donné ; en dehors des cas de supplétion marqués dans le lexique, il n'existe pas de concept équivalent à la notion de porte-manteau correspondant au regroupement de deux niveaux réalisationnels ou plus ;
  - Contrairement à PFM (Stump, 2001) où les règles porte-manteau servent à rendre compte des formes fléchies supplétives,  $\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}$  traite la supplétion par des informations morphologiques lexicales spécifiées indépendamment dans les entrées lexicales concernées ; ces informations lexicales se substituent alors pour les traits concernés par les formes ou radicaux supplétifs aux niveaux réalisationnels concernés <sup>23</sup>.

Les niveaux réalisationnels permettent ainsi de rendre compte de classes réalisationnelles de natures différentes qui ne sont pas alignées en termes d'espaces partitionnants. Ces niveaux permettent par exemple de rendre compte des classes de radicaux, thématiques et d'exponence en latin par l'existence des classes suivantes définies pour le latin à la section 1.4 du chapitre 1 :

- les classes flexionnelles A à D
- des quatre classes thématiques à thème *a*, *e*, *i* d'une part et athématique de l'autre ;
- des différentes classes de radicaux forment leur radicaux RAD2 et RAD3 par des procédés différents à chaque fois ; par exemple par reduplication ou par allongement de la voyelle ou par insertion des segments *-u-* et *-t-* respectivement, *etc.*

#### 4.1.6 Représentation des phénomènes de décalage morphosyntaxique

Le dernier élément dans la définition des entrées lexicales de  $\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}$  que nous n'avons pas encore abordé est ce que nous appelons les FONCTIONS DE TRANSFERT. Une fonction

---

23. C'est-à-dire pour les formes supplétives, à l'intégralité des niveaux réalisationnels, pour les radicaux supplétifs, aux seuls niveaux de type radical et thématique.

de transfert permet de rendre compte de phénomènes de décalage morphosyntaxique<sup>24</sup>. Les fonctions de transfert sont ainsi les pendants directs des règles de *Paradigm Linkage* en *Paradigm Function Morphology* Stump (2001, 2006) et du lien entre les chemins  $\langle \text{syn} \rangle$  et  $\langle \text{mor} \rangle$  en *Network Morphology* Brown & Hippisley (2012). Il s'agit d'une fonction définie pour une entrée lexicale donnée et pour une zone réalisationnelle donnée. Elle est comprise dans les éléments contenus dans les sous-schémas d'une entrée lexicale particulière. Un sous-schéma comporte ainsi un ensemble de couples (zone réalisationnelle, fonction de transfert) appelés COUPLES RÉALISATIONNELS. Un couple réalisationnel est habituellement noté  $p = (z, t)$ .

Une fonction de transfert prend en entrée un PTM et donne en sortie un autre PTM. Cette notion définie dans  $\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}$  permet l'encodage des cas de décalage morphosyntaxiques de la façon suivante.

Considérons les données du serbo-croate décrites au chapitre 3. Rappelons que les noms serbo-croates se fléchissent selon un certain nombre de classes flexionnelles distinctes dont celles que nous avons illustrées au tableau 3.15 page 183 selon des données issues de Baerman (2006). Nous les rappelons au tableau 4.10 en indiquant pour chaque classe les zones flexionnelles dont elles sont constituées. Certains noms comme DETE « *enfant* » et BRAT « *frère* » sont des noms qui se fléchissent respectivement selon les patrons du singulier des noms de la classe flexionnelle  $Z_{N-et}^{exp}$  des noms neutres à RADICAL EN –ET et des noms de la classe flexionnelle  $Z_M^{exp}$  des noms masculins, mais ils font partie de ces noms du serbo-croate qui emploient des formes du singulier d'une autre classe pour exprimer leur pluriel. Comme le montre le tableau 4.11, ils construisent leur pluriel en empruntant les désinences du singulier des noms féminins à RADICAL EN –A de la classe  $Z_{F-a}^{exp}$  comme ŽENA. Au delà du fait que cela fait de ces noms des noms hétéroclites, l'emploi d'un singulier pour exprimer le pluriel est une instance de décalage morphosyntaxique.

Grâce aux notions définies dans  $\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}$ , nous pouvons rendre compte explicitement et de façon séparée des deux phénomènes non canoniques à l'œuvre. Les noms sont d'une part associés (pour chacun de leur PTM) à une zone flexionnelle déterminant les règles de réalisation selon lesquelles ils se fléchissent. Pour DETE « *enfant* » et BRAT « *frère* », ces zones flexionnelles sont respectivement d'une part la zone flexionnelle du singulier  $z_{N-et,sg}^{exp}$  des noms neutres en *-et* et la zone du singulier  $z_{F-a,sg}^{exp}$  des noms féminins en *-a* pour DETE et d'autre part la zone  $z_{M-a,sg}^{exp}$  des noms masculins en *-a* et la zone du singulier  $z_{F-a,sg}^{exp}$  des noms féminins en *-a* pour BRAT. Dans la mesure où  $z_{N-et,sg}^{exp}$  et  $z_{F-a,sg}^{exp}$ , respectivement

24. Ce que Baerman (2007a) qualifie de « *déponence* » et définit comme « *un décalage entre forme et fonction* », voir chapitre 1.

	$Z_{N-et}^{exp}$ : NEUTRE EN -ET DUGME « bouton »		$Z_M^{exp}$ : MASCULIN JELEN « cerf »		$Z_{F-a}^{exp}$ : FÉMININ EN -A ŽENA « femme »	
	$z_{N-et,sg}^{exp}$ : SG	$z_{N-et,pl}^{exp}$ : PL	$z_{M,sg}^{exp}$ : SG	$z_{M,pl}^{exp}$ : PL	$z_{F-a,sg}^{exp}$ : SG	$z_{F-a,pl}^{exp}$ : PL
NOM	<i>dugme</i>	<i>dugm-et-a</i>	<i>jelen</i>	<i>jelen-i</i>	<i>žen-a</i>	<i>žen-e</i>
ACC	<i>dugme</i>	<i>dugm-et-a</i>	<i>jelen-a</i>	<i>jelen-e</i>	<i>žen-u</i>	<i>žen-e</i>
GEN	<i>dugm-et-a</i>	<i>dugm-et-a</i>	<i>jelen-a</i>	<i>jelen-a</i>	<i>žen-e</i>	<i>žen-a</i>
DAT	<i>dugm-et-u</i>	<i>dugm-et-ima</i>	<i>jelen-u</i>	<i>jelen-ima</i>	<i>žen-i</i>	<i>žen-ama</i>
LOC	<i>dugm-et-u</i>	<i>dugm-et-ima</i>	<i>jelen-u</i>	<i>jelen-ima</i>	<i>žen-i</i>	<i>žen-ama</i>
INS	<i>dugm-et-om</i>	<i>dugm-et-ima</i>	<i>jelen-om</i>	<i>jelen-ima</i>	<i>žen-om</i>	<i>žen-ama</i>

TABLE 4.10 : Déclinaison régulière de noms du serbo-croate

	NEUTRE EN -ET/ RADICAL EN -A DETE « enfant »		MASCULIN / RADICAL EN -A BRAT « frère »	
	$z_{N-et,sg}^{exp}$ : SG	$z_{F-a,sg}^{exp}$ : PL	$z_{M,sg}^{exp}$ : SG	$z_{F-a,sg}^{exp}$ : PL
NOM	<i>dete</i>	<i>dec-a</i>	<i>brat</i>	<i>brać-a</i>
ACC	<i>dete</i>	<i>dec-u</i>	<i>brat-a</i>	<i>brać-u</i>
GEN	<i>det-et-a</i>	<i>dec-e</i>	<i>brat-a</i>	<i>brać-e</i>
DAT	<i>det-et-u</i>	<i>dec-i</i>	<i>brat-u</i>	<i>brać-i</i>
LOC	<i>det-et-u</i>	<i>dec-i</i>	<i>brat-u</i>	<i>brać-i</i>
INS	<i>det-et-om</i>	<i>dec-om</i>	<i>brat-om</i>	<i>brać-om</i>

TABLE 4.11 : Décalage dans deux paradigmes de noms du serbo-croate

$z_{M-a,sg}^{exp}$  et  $z_{F-a,sg}^{exp}$ , n'appartiennent pas à la même classe flexionnelle, DETE et BRAT sont bien des hétéroclites.

Par ailleurs, la zone empruntée par chacun de ces noms à la classe flexionnelle des noms féminins en *-a* est une zone comportant des règles de réalisations pour des PTM comportant le singulier. Or, elles sont pour DETE et BRAT sélectionnées pour fournir des règles de réalisation de formes pour le pluriel. Les PTM donnés en entrée aux règles de réalisation doivent donc comporter le trait singulier. Cela est obtenu par la fonction de transfert associée aux entrées lexicales concernées pour chaque zone qui lui est associée à un niveau donné et qui transforme tous les PTM comportant le trait pluriel en un PTM identique au nombre près : le trait NOMBRE contenu dans le PTM transformé a alors pour valeur *singulier*. La réalisation de la forme peut ainsi se faire pour une forme plurielle comme si le nombre demandé était le singulier. En d'autres termes, la forme réalisée correspond à une réalisation habituelle du singulier, mais la fonction de la forme en question pour les lexèmes DETE et BRAT est d'exprimer le pluriel. Il y a donc dans ce cas un décalage morphosyntaxique. La règle de transfert  $t_{NB}$  correspondante sera notée comme

suit dans la formalisation PARSLI.

$$4.3 \quad t_{\text{NB}}(\{\text{NB } sg\}) = \{\text{NB } pl\}$$

Notons que dans le cas canonique, la fonction de transfert associée à une entrée lexicale est la fonction identité (notée *id*).

Les entrées lexicales associées à DUGME et ŽENA d’une part et DETE de l’autre sont indiquées aux figures 4.9, 4.10 et 4.11. Elles sont toutes deux des entrées nominales et comportent un inventaire standard de dix PTM, dont cinq comportant le nombre singulier et cinq le nombre pluriel. Elles sont par ailleurs associées à des schèmes. Les schèmes de DUGME et ŽENA ne comportent chacun qu’un seul sous-schème constitué d’une classe de radical  $Z_{\text{reg}}^s$  et d’une classe d’exponence  $Z_{\text{N-ET}}^{\text{exp}}$  pour les noms neutres en *-et*, respectivement  $Z_{\text{F-A}}^{\text{exp}}$  pour les noms féminins en *-a*. Le schème associé à DETE comporte deux sous-schèmes constitués de la classe de radical  $Z_{\text{reg}}^s$  régulière (sans allomorphie radicale) associée respectivement à la zone du singulier  $z_{\text{N-ET, SG}}^{\text{exp}}$  de la classe d’exponence  $Z_{\text{N-ET}}^{\text{exp}}$  et à la zone du singulier  $z_{\text{F-A, SG}}^{\text{exp}}$  de la classe d’exponence  $Z_{\text{F-A}}^{\text{exp}}$ . Chacune de ces classes ou zones est associée à sa fonction de transfert correspondante, à savoir la fonction identité pour toutes sauf  $z_{\text{F-A, SG}}^{\text{exp}}$ . Cette dernière est associée à la fonction de transfert  $t_{\text{NB}}$  permettant d’employer la zone  $z_{\text{A, SG}}^{\text{exp}}$  pour la réalisation des formes du pluriel de DETE.

DUGME	
I-PHON	dugme
I-CAT	<i>nom</i>
MSF	{ <i>standard</i> }
S-STEM	( <i>vide</i> )
S-FORM	( <i>vide</i> )
I-PAT	( $Z_{\text{reg}}^s, \text{id}$ ), ( $Z_{\text{N-ET}}^{\text{exp}}, \text{id}$ )

FIGURE 4.9 : Entrée lexicale pour le nom du serbo-croate DUGME « *bouton* »

La même chose vaut pour les entrées lexicales illustrées aux figures 4.12 et 4.13 associées à JELEN et BRAT respectivement.

ŽENA	
I-PHON	žena
I-CAT	<i>nom</i>
MSF	{ <i>standard</i> }
S-STEM	( <i>vide</i> )
S-FORM	( <i>vide</i> )
I-PAT	( $Z_{\text{reg}}^s, \text{id}$ ), ( $Z_{\text{F-A}}^{\text{exp}}, \text{id}$ )

FIGURE 4.10 : Entrée lexicale pour le nom du serbo-croate ŽENA « *femme* »

DETE	
I-PHON	dete
I-CAT	<i>nom</i>
MSF	{ <i>standard</i> }
S-STEM	( <i>vide</i> )
S-FORM	( <i>vide</i> )
I-PAT	( $Z_{\text{reg}}^s, \text{id}$ ), ( $z_{\text{N-ET, SG}}^{\text{exp}}, \text{id}$ )   ( $Z_{\text{reg}}^s, \text{id}$ ), ( $z_{\text{F-A, SG}}^{\text{exp}}, t_{\text{NB}}$ )

FIGURE 4.11 : Entrée lexicale pour le nom serbo-croate DETE « *enfant* »

JELEN	
I-PHON	jelen
I-CAT	<i>nom</i>
MSF	{ <i>standard</i> }
S-STEM	( <i>vide</i> )
S-FORM	( <i>vide</i> )
I-PAT	( $Z_{\text{reg}}^s, \text{id}$ ), ( $Z_{\text{M}}^{\text{exp}}, \text{id}$ )

FIGURE 4.12 : Entrée lexicale pour le nom du serbo-croate JELEN « *cerf* »

BRAT	
I-PHON	brat
I-CAT	<i>nom</i>
MSF	{ <i>standard</i> }
S-STEM	( <i>vide</i> )
S-FORM	( <i>vide</i> )
I-PAT	( $Z_{\text{reg}}^s, \text{id}$ ), ( $z_{M, \text{SG}}^{\text{exp}}, \text{id}$ )
	( $Z_{\text{reg}}^s, \text{id}$ ), ( $z_{F-A, \text{SG}}^{\text{exp}}, t_{\text{NB}}$ )

FIGURE 4.13 : Entrée lexicale pour le nom du serbo-croate BRAT « frère »

#### 4.1.7 Représentation de la surabondance, de la défektivité et de la surdifférenciation

Dans  $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$ , l'association, pour une entrée lexicale, d'un PTM avec une, aucune ou plusieurs zones représente le caractère respectivement canonique (à cet égard du moins), défectif ou surabondant de cette entrée pour la case du paradigme concernée par ce PTM. Dans la mesure où la réalisation des formes d'une entrée lexicale donnée passe par l'application de règles réalisationnelles provenant de zones réalisationnelles pour chaque niveau réalisationnel, la DÉFECTIVITÉ et la SURABONDANCE peuvent être définies niveau par niveau.

Nous avons donné un exemple de surabondance en français à la section 4.1.1 pour les verbes en *-ayer*. La modélisation en  $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$  exprimait cette surabondance par une association de certains PTM avec plus d'une zone de radicaux au niveau radical, cf. la figure 4.2 page 197. Cette double association avait pour effet la formation de deux radicaux distincts dans certaines cases, chacun combiné avec les exposants associés aux verbes du premier groupe, résultant ainsi en une surabondance de formes pour les cases ayant deux radicaux. Le fait que ce soit au niveau radical qu'une association surnuméraire entre PTM et zones de radicaux soit définie permet de situer l'origine de la surabondance de formes comme provenant d'une surabondance de radicaux. Cette surabondance est indiquée dans le schème des verbes en question. Le schème flexionnel  $\mathbb{P}_{\text{BALAYER}}$  d'un verbe en *-ayer* comme BALAYER peut ainsi être représenté comme indiqué ci-dessous.

$$\mathbb{P}_{\text{BALAYER}} = \left| \begin{array}{ll} (z_{ay}^s, \text{id}), & (z_{v1,1}^{exp}, \text{id}) \\ (z_{ay}^s, \text{id}), & (z_{v1,2}^{exp}, \text{id}) \\ (z_{ai}^s, \text{id}), & (z_{v1,2}^{exp}, \text{id}) \end{array} \right.$$

Dans la notation, l'ensemble correspond au schème de BALAYER et les différentes lignes correspondent aux sous-schémes respectifs. Chaque sous-schème comprend respectivement deux couples réalisationnels indiquant la zone de radical  $z^s$  ou d'exponence  $z^{exp}$  assortie d'une fonction de transfert. Dans cet exemple, la fonction de transfert est à chaque fois la fonction identité *id*.

Un autre exemple de surabondance régulière peut être observée en latin. Au parfait, les verbes des classes A et C<sup>25</sup> comportent une désinence *-ērunt*. Mais la plupart de ces verbes possèdent également une forme alternative avec la désinence *-ēre*. Ce type de surabondance sera modélisé en  $\mathbb{P}_{\text{TRAIRESL}}$  par une association du PTM 3PL.PRS.PFV.IND avec deux zones flexionnelles, l'une comportant la règle de réalisation associée à 3PL.PRS.PFV.IND consistant à ajouter le suffixe *-ērunt*, l'autre celle consistant à ajouter le suffixe *-ēre*. Dans ce cas, la surabondance est localisée dans le niveau d'exponence.

Le cas de la défektivité peut quant à lui être modélisé par l'association d'un PTM à aucune zone (à un ou plusieurs des niveaux réalisationnels) pour une entrée lexicale donnée. Un exemple de défektivité en ce sens est le verbe TRAIRE en français. Le verbe TRAIRE présente un paradigme auquel il manque les sous-paradigmes du passé-simple de l'indicatif ainsi que de l'imparfait du subjonctif. Il n'est notamment pas possible de déterminer si la forme du passé simple de l'indicatif à la première personne du singulier, si elle existait, devrait se construire comme *je \*tris* sur le modèle de FAIRE, comme *je \*trus* sur le modèle de PLAIRE ou encore, comme les verbes réguliers du premier groupe, *je \*trayai*. L'entrée lexicale de TRAIRE peut ainsi être représentée comme indiqué à la figure 4.14. Elle comporte une catégorie flexionnelle I-CAT de type *verbe*, un inventaire de PTM MSF standard et un schème flexionnel  $\mathbb{P}_{\text{TRAIRESL}}$  défektivif. En l'occurrence, les traits exprimables par TRAIRE sont les mêmes que ceux exprimés par sa catégorie, ce que nous avons rendu ici par la notation *standard*.

Le schème (simplifié ici) indique notamment que la construction des formes de TRAIRE se fait par deux niveaux réalisationnels : un niveau radical et un niveau exposant. Les sous-schémes du schème indiquent les combinaisons de zone de radical et zone flexionnelle (ou zone d'exponence) suivantes :

– Premier sous-schème :

25. Formes traditionnellement dites « *actives* », cf. chapitre 1.

- un radical TRAI obtenu par la zone de radical  $z_{ai}^s$  défini pour l'espace partitionnant comportant le présent et l'imparfait de l'indicatif et du subjonctif privés à chaque fois de la première et deuxième personnes du pluriel, ainsi que le participe passé et l'infinitif,
  - combiné à la classe flexionnelle  $Z_{\text{TRAIRE}}^{\text{exp}}$ <sup>26</sup> définie pour l'intégralité des PTM identifiés pour MSF.
- Deuxième sous-schème :
- un radical TRAY obtenu par la zone de radical  $z_{ay}^s$  défini pour l'espace partitionnant comportant la première et deuxième personnes du pluriel du présent (indicatif et subjonctif) et de l'imparfait ainsi que le participe présent,
  - combiné à la classe flexionnelle  $Z_{\text{TRAIRE}}^{\text{exp}}$  définie pour l'intégralité des PTM définis par MSF.
- Sous-schème défectif :
- les PTM comportant le passé-simple et l'imparfait du subjonctif ne sont associés à aucune zone de radical, provoquant une défectivité au niveau radical pour les sous-paradigmes concernés.

Dans la mesure où TRAIRE n'affiche pas de décalage morphosyntaxique, la fonction de transfert indiquée est à chaque fois la fonction identité id.

Le type de défectivité qu'affiche le verbe TRAIRE correspond à ce que Boyé & Cabredo Hofherr (2010) appellent de la DÉFECTIVITÉ MORPHOLOGIQUE. Elle s'oppose à un autre type de défectivité illustré notamment par le verbe SEOIR. Ce verbe ne réalise en tout que treize paquets de traits morphosyntaxiques (sur les 47 habituellement exprimées de façon synthétique par les verbes français (Arrivé, 1997)<sup>27</sup>). Le paradigme de SEOIR est ainsi réduit aux quatorze formes indiquées dans le tableau 4.12 page 222, la case du participe présent étant de surcroît surabondante.

La différence entre SEOIR et TRAIRE est que pour SEOIR, l'existence du paradigme du verbe ASSEOIR donne une idée précise de ce à quoi pourraient ressembler les autres formes de SEOIR si elles existaient. Ce type de défectivité n'est ainsi pas dû à l'inexistence de règles de réalisation appropriées pour construire les formes. En  $\mathcal{PARSL}$ , cela se traduit notamment par le fait que le schème de SEOIR est complet. La défectivité relève ici d'un

26. Pour ne pas surcharger la présentation à ce stade, nous ne rentrons pas ici dans la question du nombre de classes flexionnelles à poser pour le français. Dans notre analyse de la flexion verbale du français que nous présenterons brièvement au chapitre 6, TRAIRE est traité comme un verbe régulier du troisième groupe en –re.

27. Le Bescherelle (Arrivé, 1997) ne liste qu'une seule forme pour le participe passé.



TRAIRE	
I-PHON	traire
I-CAT	<i>verbe</i>
MSF	{ <i>standard</i> }
S-STEM	( <i>vide</i> )
S-FORM	( <i>vide</i> )
I-PAT	$(z_{ai}^s, \text{id}), (Z_{\text{TRAIRE}}^{\text{exp}}, \text{id})$ $(z_{ay}^s, \text{id}), (Z_{\text{TRAIRE}}^{\text{exp}}, \text{id})$ — $(Z_{\text{TRAIRE}}^{\text{exp}}, \text{id})$

FIGURE 4.14 : Entrée lexicale pour le verbe TRAIRE

	PRS	PST (IPRF)	FUT
3SG.IND	<i>sied</i>	<i>seyait</i>	<i>siéra</i>
3PL.IND	<i>siéent</i>	<i>seyaient</i>	<i>siéront</i>
3SG.COND	<i>siérait</i>		
3PL.COND	<i>siéraient</i>		
3SG.SBJV	<i>siée</i>		
3PL.SBJV	<i>siéent</i>		
INF	<i>seoir</i>		
PTCP	<i>séant/seyant</i>	<i>sis</i>	

TABLE 4.12 : Paradigme déficient : le verbe SEOIR

inventaire réduit de paquets de traits morphosyntaxiques par rapport à celui qui est attendu par un lexème verbal. Dans la suite de cette thèse nous parlerons de DÉFICIENCE ou DÉFECTIVITÉ MORPHOSYNTAXIQUE pour ce type de défektivité. Nous garderons le terme de DÉFECTIVITÉ MORPHOLOGIQUE proposé par Boyé & Cabredo Hofherr ou de DÉFECTIVITÉ FORMELLE pour les cas comme TRAIRE où les cases manquantes sont le résultat d'une association manquante avec une zone réalisationnelle pour un ou plusieurs PTM définis.

Nous pouvons ainsi représenter l'entrée lexicale de SEOIR comme à la figure 4.15. Elle comporte une catégorie flexionnelle I-CAT de type *verbe*, un inventaire de PTM MSF explicitement défini et déficient et un schème flexionnel  $\mathbb{P}_{\text{SEOIR}}$  surabondant. Le schème (simplifié ici) indique notamment l'emploi d'un radical SIE défini par la zone de radical  $z_{ie}^s$

par l'espace partionnant du présent et du futur avec la classe flexionnelle  $Z_{\text{SEOIR}}^{\text{exp}}$  définie pour l'intégralité des PTM définis par MSF. Les formes produites par cette association correspondent à l'intersection des espaces partionnants définissant  $z_{ie}^s$  et  $Z_{\text{SEOIR}}^{\text{exp}}$ . Les autres associations sont celles de  $Z_{\text{SEOIR}}^{\text{exp}}$  avec respectivement  $z_{ey}^s$ , pour le radical SEY définie par l'espace couvrant le passé et le participe et  $z_e^s$  pour le radical SE définie par l'espace couvrant l'infinitif et le participe. Dans la mesure où SEOIR n'affiche pas de décalage morphosyntaxique, la fonction de transfert indiquée est à chaque fois la fonction identité  $\text{id}$ <sup>28</sup>.

SEOIR	
I-PHON	seoir
I-CAT	verbe
MSF	$\left\{ \begin{array}{l} 3\text{SG.PRS.IND}, 3\text{PL.PRS.IND}, 3\text{SG.PST.IND}, \\ 3\text{PL.PST.IND}, 3\text{SG.FUT.IND}, 3\text{PL.FUT.IND}, \\ 3\text{SG.PRS.COND}, 3\text{PL.PRS.COND}, 3\text{SG.PRS.SBJV}, \\ 3\text{PL.PRS.SBJV}, \text{INF}, \text{PTCP} \end{array} \right\}$
S-STEM	(vide)
S-FORM	(vide)
I-PAT	$(z_{ie}^s, \text{id}), (Z_{\text{SEOIR}}^{\text{exp}}, \text{id})$ $(z_{ey}^s, \text{id}), (Z_{\text{SEOIR}}^{\text{exp}}, \text{id})$ $(z_e^s, \text{id}), (Z_{\text{SEOIR}}^{\text{exp}}, \text{id})$

FIGURE 4.15 : Entrée lexicale pour le verbe SEOIR

Le cas de la SURDIFFÉRENCIATION apparaît enfin comme le cas inverse de la déficience ou défektivité morphosyntaxique. Dans PARSLI, la surdifférenciation sera modélisée comme mettant en jeu, pour un lexème donné, un inventaire de PTM plus étendu que celui habituellement réalisé par les lexèmes de la catégorie à laquelle il appartient. Nous donnons au tableau 4.13 l'exemple du nom slovaque PRIATEL « *ami* » (Zauner, 1973). Contrairement aux autres noms du slovaque qui comme BÁDATEL « *chercheur scientifique* » ont un paradigme à six cas, ce nom comporte, comme quelques rares autres

28. La distinction entre défektivité morphologique et défektivité morphosyntaxique est orthogonale à celle faite par Boyé & Cabredo Hofherr (2010) entre défektivité morphologique et défektivité syntaxique, qui est basée sur l'existence ou non d'une motivation extramorphologique de la défektivité. Dans nos termes, le verbe FALLOIR est déficient, comme le verbe SEOIR, bien que dans son cas la déficience ait une motivation syntaxique et sémantique qui est absente pour SEOIR.

noms<sup>29</sup>, un cas vocatif surnuméraire au singulier.

	BÁDATEL « chercheur »		PRIATEL « ami »	
	SINGULIER	PLURIEL	SINGULIER	PLURIEL
NOM	<i>bádatel'</i>	<i>bádatelia</i>	<i>priatel'</i>	<i>priatelia</i>
GEN	<i>bádatela</i>	<i>bádatelov</i>	<i>priatela</i>	<i>priatelov</i>
DAT	<i>bádatelovi</i>	<i>bádatelom</i>	<i>priatelovi</i>	<i>priatelom</i>
ACC	<i>bádatela</i>	<i>bádatelov</i>	<i>priatela</i>	<i>priatelov</i>
VOC	–	–	<i>priatel'u</i>	–
LOC	<i>bádatelovi</i>	<i>bádateloch</i>	<i>priatelovi</i>	<i>priateloch</i>
INS	<i>bádatelom</i>	<i>bádatel'mi</i>	<i>priatelom</i>	<i>priatel'mi</i>

TABLE 4.13 : Paradigmes masculins animés consonantiques en slovaque, dont l'un manifeste de la surdifférenciation au VOC.SG (données de (Zauner, 1973)).

L'entrée lexicale nominale (simplifiée) de PRIATEL, illustrée à la figure 4.16, comporte donc treize PTM au lieu de douze, ainsi qu'un schème avec deux sous-schémas dont un avec une zone flexionnelle du vocatif  $z_{\text{VOC}}^{\text{exp}}$  spécifique aux noms de ce type. La zone de radical  $z_{\text{C}}^{\text{s}}$  indiquée est par ailleurs celle des noms à radical consonantique ; la classe flexionnelle  $Z_{\text{M.ANIM-C}}^{\text{exp}}$  correspond à une classe flexionnelle pour certains noms masculins animés à radical consonantique.

#### 4.1.8 Résumé : modélisation de dix phénomènes non canoniques en $\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}\mathbb{I}$

Dans cette section nous avons évoqué neuf des dix phénomènes non canoniques dont permet de rendre compte  $\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}\mathbb{I}$ . La SUPPLÉTION DE RADICAL, la SUPPLÉTION DE FORME, la DÉFECTIVITÉ MORPHOLOGIQUE (OU FORMELLE), la SURABONDANCE, l'HÉTÉROCLISE, la SURDIFFÉRENCIATION et le DÉCALAGE (MORPHOSYNTAXIQUE) avaient déjà été décrits au chapitre 3. Nous avons également montré qu'il est possible de différencier la supplétion de radical au sens strict de la simple ALLOMORPHIE RADICALE régulière. Nous avons de surcroît défini le phénomène de DÉFICIENCE (OU DÉFECTIVITÉ MORPHOSYNTAXIQUE) qui se distingue de la défektivité formelle par les facteurs étant à l'origine des cases absentes du paradigmes : dans le cas de la défektivité, une absence de zone réalisationnelle associée à certains PTM (à au moins un niveau réalisationnel) ; un

29. Zauner (1973) indique notamment les lexèmes suivants : SYN « fils » (vocatif *synku*), BRAT « frère » (vocatif *brat(k)u*), CHLAPEC « garçon » (vocatif *chlapče*), ŠVAGOR « beau-frère » (vocatif *švagre*), KMOTOR « parrain » (vocatif *kmotre*), CHLAP « homme » (vocatif *chlape*), PÁN « Monsieur » (vocatif *pane*), MAJSTER « maître » (vocatif *majtre*), BOH « dieu » (vocatif *bože*), ČLOVEK « individu » (vocatif *človece*).

PRIATEL	
I-PHON	priateľ
I-CAT	<i>nom</i>
MSF	$\left\{ \begin{array}{l} \text{NOM.SG, GEN.SG, DAT.SG, ACC.SG,} \\ \text{VOC.SG, LOC.SG, INS.SG,} \\ \text{NOM.PL, GEN.PL, DAT.PL, ACC.PL,} \\ \text{LOC.PL, INS.PL} \end{array} \right\}$
S-STEM	( <i>vide</i> )
S-FORM	( <i>vide</i> )
I-PAT	$\begin{array}{l} (z_C^s, \text{id}), (Z_{M.ANIM-C}^{exp}, \text{id}) \\ (z_C^s, \text{id}), (z_{VOC}^{exp}, \text{id}) \end{array}$

FIGURE 4.16 : Entrée lexicale pour le nom slovaque PRIATEL

inventaire de PTM réduits dans le cas de la déficience. À ces neuf phénomènes s’ajoute enfin le SYNCRÉTISME représenté dans  $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$  par une simple identité des formes produites.

Nous résumons, à la figure 4.17, les différents phénomènes non canoniques modélisables en  $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$ . Cette figure fait écho à la figure 3.10 du chapitre 3 en y intégrant le phénomène de déficience nouvellement défini<sup>30</sup>. Nous donnons également les propriétés de modélisation de chacun de ces phénomènes dans le tableau 4.14 faisant écho au tableau 3.17 de définition des phénomènes non canoniques du chapitre 3. L’annexe C comporte enfin une discussion de l’interaction des notions prise en compte par  $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$  avec une autre dimension de la canonicité développée par Corbett (2010), à savoir une théorie indépendante de la notion de hiérarchie de traits morphosyntaxiques. Cette dimension est cependant orthogonale aux questions de canonicité de la structure paradigmatique et n’est à ce titre pas directement intégré dans le modèle  $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$ .

30. Nous avons étiqueté chacune des branches par une lettre (F, C, T) correspondant au type d’irrégularités observables pour chacun des phénomènes : F pour des irrégularités portant sur la réalisation des formes du paradigme, C pour les irrégularités dans le remplissage des cases du paradigme et T pour les irrégularités concernant la réalisation des traits morphosyntaxiques.

FIGURE 4.17 : Axes de non-canonicité dans  $\mathbb{P}^n$

	PHÉNOMÈNE	DESCRIPTION	PARSLI
F1	SUPLÉTION DE RADICAL	<i>Le paradigme affiche des alternances de radical qui ne sont pas construites régulièrement.</i>	<i>Un ou plusieurs radicaux supplétifs sont listés dans l'entrée lexicale.</i>
F2	SUPLÉTION DE FORME	<i>Le paradigme comporte des formes qui remplacent à la fois le radical et les exposants attendus pour une case donnée.</i>	<i>Une ou plusieurs formes supplétives sont listées dans l'entrée lexicale.</i>
F3	SYNCRÉTISME	<i>Le paradigme comporte deux formes identiques.</i>	<i>(Les formes produites sont identiques.)</i>
C1	DÉFECTIVITÉ FORMELLE	<i>Le paradigme comporte des cases vides.</i>	<i>À au moins un niveau réalisationnel, au moins un PTM n'est pas associé à une zone réalisationnelle.</i>
C2	SURABONDANCE	<i>Le paradigme comporte des cases avec plus d'une forme.</i>	<i>À au moins un niveau réalisationnel, au moins un PTM est associé à plus d'une zone réalisationnelle.</i>
C3	ALLORMORPHIE RADICALE	<i>Le paradigme affiche des alternances construites régulièrement.</i>	<i>À au moins un niveau réalisationnel radical, la construction du radical n'est pas prise en charge par une zone réalisationnelle unique.</i>
C4	HÉTÉROCLISE	<i>Les formes du paradigmes sont construites à partir d'au moins deux classes flexionnelles distinctes.</i>	<i>À un niveau réalisationnel, les zones employées pour la construction des formes appartiennent à des classes flexionnelles différentes.</i>
T1	DÉFICIENCE (DÉFECTIVITÉ MORPHOSYNTAXIQUE)	<i>Il manque des cases au paradigme.</i>	<i>L'entrée lexicale est définie pour un nombre de PTM inférieur à celui de sa catégorie flexionnelle.</i>
T2	SURDIFFÉRENCIATION	<i>Le paradigme comporte un surplus de cases dû à la distinction de traits ou des valeurs de traits supplémentaires.</i>	<i>L'entrée lexicale est définie pour un nombre de PTM supérieur à celui de sa catégorie flexionnelle.</i>
T3	DÉCALAGE	<i>Un ou plus d'un PTM sont exprimés par une forme attendue pour un autre PTM.</i>	<i>Dans le schème, au moins une fonction de transfert n'est pas la fonction identité id.</i>

TABLE 4.14 : Récapitulatif des neuf types de phénomènes non canoniques dans PARSLI

## 4.2 Présentation des primitives du modèle

Dans cette section, nous présentons de façon plus formalisée les primitives du modèle morphologique flexionnel  $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$  dont nous avons indiqué les caractéristiques principales à la section précédente.

Nous organisons cette présentation en prenant pour point de départ la définition d'une ENTRÉE LEXICALE dans  $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$ . En particulier, la définition formelle de la notion de SCHÈME FLEXIONNEL, et par ce biais celle des notions de SOUS-SCHÈME, de COUPLE RÉALISATIONNEL, de ZONE RÉALISATIONNELLE et de FONCTION DE TRANSFERT, nous conduiront naturellement à décrire le MÉCANISME DE RÉALISATION des FORMES dans  $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$ . Ces notions font partie des primitives du modèle.

Mais l'objectif de  $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$  ne se résume cependant pas à la définition de ce qui constitue une entrée lexicale et à la description de la réalisation flexionnelle. Il s'agit également de formuler une THÉORIE des éléments constitutifs de la GRAMMAIRE FLEXIONNELLE comme les CLASSES FLEXIONNELLES ou du lexique comme les CATÉGORIES FLEXIONNELLES.

La dimension flexionnelle des modules grammatical et lexical de  $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$  est conçue de telle sorte que ses caractéristiques sont déductibles des généralisations observables dans les paradigmes des flexèmes de la langue. Les COMPORTEMENTS MAJORITAIRES OBSERVABLES dans ces paradigmes flexionnels correspondent ainsi à une régularité du système. Nous les traitons comme le comportement par défaut de la langue en question.

Plus précisément, ces comportements par défaut déterminent la définition des catégories flexionnelles, comme nous allons le voir ci-dessous. Ils déterminent également au sein de la grammaire la définition des classes flexionnelles, des thèmes et des classes de radicaux, comme nous le verrons à la fin de cette section.

Les symboles usuellement employées en  $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$  pour désigner les éléments de la grammaire flexionnelle sont inventoriés à la figure 4.18.

### 4.2.1 Entrées lexicales flexionnelles

$\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$  repose sur une définition explicite des éléments constitutifs d'une ENTRÉE LEXICALE MORPHOLOGIQUE FLEXIONNELLE. La définition d'une entrée lexicale dans  $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$  se résume à la définition des informations intervenant dans la réalisation flexionnelle telles qu'elles se reflètent dans la structure du paradigme flexionnel qui lui est associé. En particulier,  $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$  ne distingue pas deux entrées lexicales qui, quoique comprenant un sens lexical différent ou étant associées à des structures syntaxiques différentes, sont parfaitement identiques dans la réalisation de leurs paradigmes flexionnels

1. Un système morphologique flexionnel :  $\mu_{\text{infl}}$
2. Les traits morphosyntaxiques :
  - Un ensemble de paquets de traits morphosyntaxiques (PTM) :  $\Sigma$
  - Un paquet de traits morphosyntaxiques particulier :  $\sigma$
3. Le lexique :
  - Une catégorie flexionnelle :  $\mathcal{C}$
  - Un ensemble de catégories flexionnelles :  $\mathcal{C}$
  - Le lexique flexionnel :  $\mathcal{I}$
  - Une entrée lexicale :  $i$
  - La base phonologique donnée en entrée aux règles de flexion :  $b$
  - Un ensemble de radicaux supplétifs :  $S_i^{\text{suppl}}$
  - Un ensemble de formes supplétives :  $F_i^{\text{suppl}}$
  - Un schème flexionnel :  $\mathbb{P}$
  - Un sous-schème :  $\mathbb{p}$
  - Un couple réalisationnel :  $p = (z, t)$
  - Une zone réalisationnelle :  $z$
  - Une fonction de transfert :  $t$
4. La grammaire
  - Un espace partitionnant :  $\pi$
  - Une zone/classe de radicaux :  $z^s/Z^s$
  - Une zone/classe thématique :  $z^h/Z^h$
  - Une zone/classe d'exponence/flexionnelle :  $z^{\text{exp}}/Z^{\text{exp}}$
  - Une zone/classe réalisationnelle :  $z/Z$
  - Un niveau réalisationnel :  $l$
  - Un ensemble de  $n$  niveaux réalisationnels :  $\{l_1, \dots, l_n\}$
  - Une règle de sélection de couples réalisationnels :  $\mathbb{L}_l$
  - Un ensemble de règles de sélection de couples réalisationnels :  $\{\mathbb{L}_{l_1}, \dots, \mathbb{L}_{l_n}\}$
  - Une règle de sélection de zones réalisationnelles :  $\mathbb{L}_l|Z$
  - Un ensemble de règles de sélection de zones réalisationnelles :  $\{\mathbb{L}_{l_1}|Z, \dots, \mathbb{L}_{l_n}|Z\}$
  - Une règle de réalisation :  $R$
  - Le résultat d'une règle de réalisation :  $r$
5. Les objets construits
  - Un radical :  $s$
  - Une forme :  $f$
  - Un paradigme :  $P$
  - Une case dans un paradigme :  $C$

FIGURE 4.18 : Notation des éléments de la grammaire dans  $\mathcal{PARSL}$



respectifs.

Les entrées lexicales définies dans  $\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}$  représentent donc des FLEXÈMES au sens de Fradin & Kerleroux (2003). Pour reprendre un exemple de Fradin & Kerleroux, la notion de flexème permet d'exprimer le fait que deux lexèmes homonymes comme  $\text{VOLER}_1$  et  $\text{VOLER}_2$  affichent en tant que flexème  $\text{VOLER}$  le même paradigme flexionnel quand bien même ils diffèrent en tant que lexèmes par les caractéristiques suivantes :

1. Leur sens :
  - $\text{VOLER}_1$  : « *se déplacer dans les airs* »
  - $\text{VOLER}_2$  : « *dérober* »
2. Leur structure argumentale :
  - $\text{VOLER}_1$  : *intransitif*
  - $\text{VOLER}_2$  : *transitif*
3. Leurs propriétés dérivationnelles :
  - $\text{VOLER}_1$  : ne permet pas de construction d'un nom d'agent en *-eur*
  - $\text{VOLER}_2$  : permet la construction du nom d'agent  $\text{VOLEUR}$  « *personne qui dérobe* ».

Dans  $\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}$ , une entrée lexicale notée  $i$  est définie par le sextuplet  $(b_i, \Sigma_i, \mathcal{C}_i, \mathbb{P}_i, S_i^{\text{suppl}}, F_i^{\text{suppl}})$  défini ci-dessous.

1. Une séquence phonologique originale, ou base phonologique, servant d'entrée aux règles de réalisation de  $i$ , notée  $b_i$ ,
2. l'ENSEMBLE DE PAQUETS DE TRAITs MORPHOSYNTAXIQUES  $\Sigma_i$  pour lesquels  $i$  se fléchit,
3. une CATÉGORIE FLEXIONNELLE notée  $\mathcal{C}_i$ ,
4. un SCHÈME noté  $\mathbb{P}_i$  qui définit la flexion de  $i$ ;
5. un inventaire noté  $S_i^{\text{suppl}}$  de RADICAUX LEXICALEMENT SPÉCIFIÉS éventuellement réduit à l'ensemble vide,
6. un inventaire noté  $F_i^{\text{suppl}}$  de FORMES LEXICALEMENT SPÉCIFIÉES éventuellement réduit à l'ensemble vide.

La figure 4.19 représente une entrée lexicale type dans  $\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}$  avec les notations employées pour représenter ses éléments constitutifs. Dans la suite de cette section nous détaillons chacun de ces éléments, leur rôle dans la réalisation des formes et leur interaction avec la canonicité du paradigme d'une entrée donnée.

LÉGENDE :

$i$	une entrée lexicale dans une lexique flexionnel $\mathcal{I}$	
I-PHON	une séquence phonologique originale servant d'entrée aux règles de réalisation	
I-CAT	la catégorie flexionnelle de $i$ notée $\mathcal{C}_i$ appartenant à l'ensemble $\mathcal{C}$ des catégories flexionnelles d'une langue	
MSF	un ensemble de traits morphosyntaxiques exprimés par $i$ et noté $\Sigma_i$	
$\sigma$	un paquet de traits morphosyntaxiques (PTM)	
S-STEM	une ensemble canoniquement vide de radicaux supplétifs, un radical noté $s$ étant constitué de sa zone de radical $z^s$ et de sa réalisation formelle $r^s$	
S-FORM	une ensemble canoniquement vide de formes supplétives, une forme $f$ étant constituée du PTM $\sigma$ qu'elle exprime et de sa réalisation formelle $r^f$	
I-PAT	le schème flexionnel de $i$ comprenant un ensemble de sous-schémas, noté $\mathbb{P}_i$	
$\mathbb{P}$	un sous-schème flexionnel comportant des couples réalisationnels notés $p$ de forme $(z, t)$	
$z$	une zone réalisationnelle	
$t_z$	la fonction de transfert associée à $z$ dans un couple réalisationnel $p$	

$i$	<table border="0"> <tr> <td style="padding-right: 10px;">I-PHON</td> <td><math>b_i</math></td> </tr> <tr> <td style="padding-right: 10px;">I-CAT</td> <td><math>\mathcal{C}_i</math></td> </tr> <tr> <td style="padding-right: 10px;">MSF</td> <td><math>\Sigma_i = \{\sigma_1, \dots, \sigma_n\}</math></td> </tr> <tr> <td style="padding-right: 10px;">S-STEM</td> <td><math>\{(z_1^s, r_1^s), \dots, (z_n^s, r_n^s)\}</math></td> </tr> <tr> <td style="padding-right: 10px;">S-FORM</td> <td><math>\{(\sigma_1, r_1^f), \dots, (\sigma_n, r_n^f)\}</math></td> </tr> <tr> <td style="padding-right: 10px;">I-PAT</td> <td> <math>\mathbb{P}_1 = p_{1,1}, \dots, p_{1,m}</math>  <math>\vdots</math>  <math>\mathbb{P}_n = p_{n,1}, \dots, p_{n,m}</math> </td> </tr> </table>	I-PHON	$b_i$	I-CAT	$\mathcal{C}_i$	MSF	$\Sigma_i = \{\sigma_1, \dots, \sigma_n\}$	S-STEM	$\{(z_1^s, r_1^s), \dots, (z_n^s, r_n^s)\}$	S-FORM	$\{(\sigma_1, r_1^f), \dots, (\sigma_n, r_n^f)\}$	I-PAT	$\mathbb{P}_1 = p_{1,1}, \dots, p_{1,m}$ $\vdots$ $\mathbb{P}_n = p_{n,1}, \dots, p_{n,m}$
I-PHON	$b_i$												
I-CAT	$\mathcal{C}_i$												
MSF	$\Sigma_i = \{\sigma_1, \dots, \sigma_n\}$												
S-STEM	$\{(z_1^s, r_1^s), \dots, (z_n^s, r_n^s)\}$												
S-FORM	$\{(\sigma_1, r_1^f), \dots, (\sigma_n, r_n^f)\}$												
I-PAT	$\mathbb{P}_1 = p_{1,1}, \dots, p_{1,m}$ $\vdots$ $\mathbb{P}_n = p_{n,1}, \dots, p_{n,m}$												

FIGURE 4.19 : Notation pour une entrée lexicale standard dans PARSLI.

À partir d'une telle entrée lexicale,  $\mathbb{P}\mathbb{R}\mathbb{S}\mathbb{L}$  définit la façon dont sont réalisées les formes qui constituent le paradigme d'un flexème  $i$ . Le PARADIGME D'UN FLEXÈME  $i$  est ainsi défini comme étant l'ensemble des formes que l'on peut construire à partir de la définition de  $i$ , chaque FORME étant un couple de la forme  $(\sigma, f)$  où  $\sigma$  est un PTM et  $f$  une réalisation constituant une forme fléchie, comme nous l'avons vu à la section 4.1.5.2. Si l'on n'applique que les niveaux radicaux et thématiques, le résultat est un RADICAL, défini comme étant un couple de la forme  $(z^s, s)$ , où  $z^s$  est une ZONE de radical et  $s$  une réalisation constituant un radical complet au sens de Matthews (1974).

### 4.2.2 Les traits morphosyntaxiques

Un PTM  $\sigma$  correspond à l'ensemble des traits exprimés par une case  $C$  du paradigme  $P$  d'une entrée lexicale  $i$  donnée. L'ensemble  $\Sigma_i$  des PTM définissant l'entrée lexicale  $i$  correspond ainsi à l'ensemble des PTM qu'exprime l'ensemble des cases du paradigme en question, autrement dit l'ensemble de PTM pour lesquels un flexème se fléchit. Dans le cas canonique, cet ensemble est le même pour tous les flexèmes appartenant à la même catégorie flexionnelle. Toutefois, dans les cas de déficience ou de surdifférenciation, l'ensemble  $\Sigma_i$  exprimé par un flexème  $i$  est différent de celui généralement exprimé par la catégorie flexionnelle  $\mathcal{C}_i$  de ce flexème.

### 4.2.3 Les catégories flexionnelles

De façon informelle, une CATÉGORIE FLEXIONNELLE  $\mathcal{C}$  est un ensemble de flexèmes partageant un nombre significatif de propriétés flexionnelles<sup>31</sup>. Une catégorie flexionnelle se distingue de la notion plus commune de catégorie morphosyntaxique par le fait qu'elle ne prend pas en compte le statut syntaxique des formes d'une entrée lexicale. Elle recouvre une réalité strictement flexionnelle qui correspond le plus souvent à des réalités syntaxiques — sans que ce ne soit nécessairement le cas. Cette réalité syntaxique en tant telle n'est cependant pas exprimée au niveau flexionnel du lexique morphologique. Elle sera encodée dans un niveau lexical supérieur prenant également en compte des informations sémantiques et syntaxiques d'une entrée donnée, c'est-à-dire d'un lexème au sens de Fradin & Kerleroux (2003).

Dans le cas canonique, une entrée  $i = (b_i, \Sigma_i, \mathcal{C}_i, \mathbb{P}_i, S_i^{suppl}, F_i^{suppl})$  se fléchit pour le même ensemble de PTM que sa catégorie flexionnelle :  $\Sigma_i = \Sigma_{\mathcal{C}_i}$ . Toujours dans le cas canonique,

---

31. En particulier les formes de leurs paradigmes expriment généralement le même ensemble de PTM réalisés par les mêmes règles de réalisation.

les entrées appartenant à une même catégorie flexionnelle réalisent également les formes de leur paradigme selon les mêmes zones réalisationnelles. Les flexèmes qui permettent ainsi la définition d'une catégorie seront considérés comme les membres canoniques de cette catégorie flexionnelle. La définition d'une catégorie flexionnelle  $\mathcal{C}$  est ainsi dérivée de la similitude significative de comportement d'un ensemble de flexèmes donnés. En ce sens, la notion de catégorie morphologique ne constitue pas une notion primitive dans le modèle PARSLI.

Un exemple extrême de catégorie flexionnelle est la catégorie des flexèmes invariables. Tout flexème  $i$  appartenant (canoniquement) à la catégorie  $\mathcal{C}_{inv}$  des mots invariables est défini par un ensemble de PTM  $\Sigma_i$  comportant un seul PTM de type « *pas de trait* » par convention noté  $\sigma_\varepsilon$ . Ce sont des flexèmes qui sont associés à un sous-schème unique comportant lui-même un seul couple réalisationnel de type radical à une seule règle réalisationnelle  $id_\varepsilon$  réalisant comme unique forme du paradigme de  $i$  une forme identique à sa forme de citation.

Par ailleurs, certains flexèmes partagent avec les flexèmes d'une catégorie flexionnelle donnée une majorité de leurs PTM, mais pas tous ; ils réalisent souvent également leurs formes selon le même ensemble de couples réalisationnels. Ils sont alors considérés comme appartenant à la catégorie flexionnelle dont les propriétés en termes de PTM et d'association avec des zones sont les plus proches de celles qu'ils affichent, mais ils en sont des membres non canoniques. C'est le cas des flexèmes déficients ou présentant des cas de surdifférentiation : pour ces flexèmes, l'ensemble des PTM pour lequel un flexème est défini peut différer de celui par lequel est défini sa catégorie flexionnelle. C'est par exemple le cas de du flexème français VIVRES dont l'entrée lexicale est donnée à la figure 4.22, à comparer avec celle du flexème LIVRE donnée à la figure 4.21. La seconde se fléchit (à l'écrit) selon les deux cases habituelles des noms du français, alors que la première ne réalise que le pluriel. On note cependant que ces deux entrées sont associées aux classes réalisationnelles  $Z_{reg}^s$  et  $Z_{reg}^{exp}$  régulières pour les noms du français : c'est notamment ce qui motive le rattachement du lexème VIVRES à la même catégorie que celle du lexème LIVRE, quoique de façon non canonique. La flexion régulière des noms du français est indiquée à la figure 4.20. Tout flexème se comportant de façon canonique selon sa catégorie indique dans son entrée lexicale qu'il possède un inventaire *standard* de PTM, comme c'est le cas pour le nom LIVRE à la figure 4.21, et qu'il se fléchit selon les couples réalisationnels *standard* de sa catégorie flexionnelle<sup>32</sup>.

32. Cette notation a pour but de montrer de façon plus immédiate l'appartenance plus ou moins canonique d'un flexème à sa catégorie flexionnelle. Rien n'empêche cependant d'explicitier pour chaque flexème les PTM pour lesquels il se fléchit ou les sous-schèmes constituant son schème flexionnel.

NOM	
I-CAT	<i>nom</i>
MSF	{ SG, PL }
I-PAT	( $Z_{\text{reg}}^s, \text{id}$ ), ( $Z_{\text{reg}}^{\text{exp}}, \text{id}$ )

FIGURE 4.20 : Catégorie flexionnelle nominale en français

LIVRE	
I-PHON	livre
I-CAT	<i>nom</i>
MSF	{ <i>standard</i> }
S-STEM	( <i>vide</i> )
S-FORM	( <i>vide</i> )
I-PAT	<i>standard</i>

FIGURE 4.21 : Entrée lexicale pour le nom LIVRE

VIVRES	
I-PHON	vivres
I-CAT	<i>nom</i>
MSF	{ <b>PL</b> }
S-STEM	( <i>vide</i> )
S-FORM	( <i>vide</i> )
I-PAT	<i>standard</i>

FIGURE 4.22 : Entrée lexicale pour le nom VIVRES

Formellement, une entrée donnée appartient à une catégorie flexionnelle si et seulement si sa valeur I-CAT est unifiable avec la valeur I-CAT de la catégorie en question. Elle en est un membre canonique si et seulement si son inventaire de PTM MSF et son schème flexionnel I-PAT sont *standard*.

#### 4.2.4 Les schèmes flexionnels

L'élément le plus complexe dans une entrée lexicale  $i$  est son schème flexionnel. C'est lui qui spécifie la façon dont sont construites les formes du paradigme de  $i$ , par l'application successive de règles de réalisation des différents niveaux réalisationnels impliqués. En particulier, c'est le schème qui spécifie quelles sont les combinaisons licites de règles de réalisation.

En première approximation, un SCHÈME  $\mathbb{P}$  peut être vu comme un ensemble de SOUS-SCHÈMES, chacun d'entre eux étant constitué d'un vecteur de COUPLES RÉALISATIONNELS qui spécifient la façon dont le niveau  $l$  intervient pour construire les formes de  $i$  produites *via* le sous-schème correspondant. Ainsi, un schème  $\mathbb{P}$  comportant  $n$  sous-schèmes et impliquant  $m$  niveaux de réalisation est une matrice de taille  $n \times m$  dont l'élément  $\mathbb{P}_{j,k}$  est un couple réalisationnel de la forme  $(z_{j,k}, t_{j,k})$ , où  $z_{j,k}$  est la ZONE RÉALISATIONNELLE de niveau réalisationnel  $k$  utilisée par le  $j$ -ième sous-schème, et où  $t_{j,k}$  est la FONCTION DE TRANSFERT correspondante. Nous allons tout d'abord définir puis illustrer sur un cas particulier la notion de zone réalisationnelle, dans le cas canonique où toutes les fonctions de transfert sont égales à la fonction identité, cas où elles ne jouent aucun rôle. Nous reviendrons ensuite sur ces fonctions de transfert.

$$\mathbb{P} = \begin{pmatrix} (z_{1,1}, t_{1,1}), & (z_{1,2}, t_{1,2}), & (z_{1,3}, t_{1,3}), & \dots, & (z_{1,m}, t_{1,m}) \\ (z_{2,1}, t_{2,1}), & (z_{2,2}, t_{2,2}), & (z_{2,3}, t_{2,3}), & \dots, & (z_{2,m}, t_{2,m}) \\ (z_{3,1}, t_{3,1}), & (z_{3,2}, t_{3,2}), & (z_{3,3}, t_{3,3}), & \dots, & (z_{3,m}, t_{3,m}) \\ \vdots & & & & \\ (z_{n,1}, t_{n,1}), & (z_{n,2}, t_{n,2}), & (z_{n,3}, t_{n,3}), & \dots, & (z_{n,m}, t_{n,m}) \end{pmatrix}$$

##### 4.2.4.1 Zones réalisationnelles et espaces partitionnants

Une zone réalisationnelle est définie comme étant une fonction partielle associant à certains PTM une règle réalisation. Dans un souci de lisibilité, nous dénoterons une zone de niveau radical au moyen du symbole  $z^s$ , une zone de niveau thématique au moyen du symbole  $z^{th}$  et une zone de niveau exposant par le symbole  $z^{exp}$ . Considérons le cas

canonique où toutes les fonctions de transfert sont égales à la fonction identité  $\text{id}$  (et ne jouent donc aucun rôle).

Considérons un sous-schème  $\mathbb{P}$  défini par le schème  $\mathbb{P}$ . Le couple réalisation unique spécifié par  $\mathbb{P}$  pour un niveau  $l$  donné associe des règles de réalisation à un certain ensemble de PTM. Le sous-schème  $\mathbb{P}$  permet ainsi de construire une forme complète pour un ensemble de PTM qui est l'intersection des ensembles de PTM couverts à chaque niveau. Ces différents ensembles de PTM sont appelés ESPACES PARTITIONNANTS. La figure 4.23 montre un exemple de la combinaison de trois espaces partitionnants  $\pi_1$ ,  $\pi_2$  et  $\pi_3$  définissant les zones réalisationnelles d'un sous-schème à trois niveaux réalisational 1, 2 et 3. Ces espaces partitionnants sont respectivement :

$$\pi_1 = \{\sigma_2, \sigma_3, \sigma_4, \sigma_5\}$$

$$\pi_2 = \{\sigma_3, \sigma_4, \sigma_5, \sigma_6, \sigma_7, \sigma_8, \sigma_9\}$$

$$\pi_3 = \{\sigma_1, \sigma_3, \sigma_5, \sigma_6, \sigma_8, \sigma_{10}\}$$

Les cases réalisées par ce sous-schème sont celles définies par l'intersection de  $\pi_1$ ,  $\pi_2$  et  $\pi_3$  :

$$\pi_1 \cap \pi_2 \cap \pi_3 = \{\sigma_3, \sigma_5\}$$

niveau 1		niveau 2		niveau 3		combinaison des niveaux	
$\sigma_1$	$\sigma_6$	$\sigma_1$	$\sigma_6$	$\sigma_1$	$\sigma_6$	$\sigma_1$	$\sigma_6$
$\sigma_2$	$\sigma_7$	$\sigma_2$	$\sigma_7$	$\sigma_2$	$\sigma_7$	$\sigma_2$	$\sigma_7$
$\sigma_3$	$\sigma_8$	$\sigma_3$	$\sigma_8$	$\sigma_3$	$\sigma_8$	$\sigma_3$	$\sigma_8$
$\sigma_4$	$\sigma_9$	$\sigma_4$	$\sigma_9$	$\sigma_4$	$\sigma_9$	$\sigma_4$	$\sigma_9$
$\sigma_5$	$\sigma_{10}$	$\sigma_5$	$\sigma_{10}$	$\sigma_5$	$\sigma_{10}$	$\sigma_5$	$\sigma_{10}$
$\pi_1 = \{\sigma_2, \sigma_3, \sigma_4, \sigma_5\}$		$\pi_2 = \{\sigma_3, \sigma_4, \sigma_5, \sigma_6, \sigma_7, \sigma_8, \sigma_9\}$		$\pi_3 = \{\sigma_1, \sigma_3, \sigma_5, \sigma_6, \sigma_8, \sigma_{10}\}$		$\pi_1 \cap \pi_2 \cap \pi_3 = \{\sigma_3, \sigma_5\}$	

FIGURE 4.23 : Combinaison des espaces partitionnants inter-niveaux

Considérons à nouveau l'entrée lexicale  $i = (b_i, \Sigma_i, \mathcal{C}_i, \mathbb{P}_i, S_i^{suppl}, F_i^{suppl})$ . Pour un certain PTM  $\sigma \in \Sigma_i$  auquel un certain sous-schème  $\mathbb{P}$  défini par  $\mathbb{P}_i$  associe des règles de réalisation à chaque niveau, la forme correspondant à  $\sigma$  et produite par  $\mathbb{P}$  s'obtient par l'application successive des couples réalisationnels inclus dans  $\mathbb{P}$ . Soit  $p = (z, \text{id})$  celui de niveau  $l$ , où  $z$  est une zone réalisationnelle. Appliquer le couple réalisationnel  $p$  revient alors à appliquer

la zone  $z$  au PTM  $\sigma$ . La règle de réalisation associée par  $\mathbb{P}$  au niveau  $l$  est alors simplement la règle de réalisation  $z(\sigma)$ . Dans le cas général,  $z(\sigma)$  prend en entrée le résultat de l'application des règles de réalisation des niveaux 1 à  $l - 1$ , et qui donne en sortie ou bien l'entrée de la règle de réalisation de niveau  $l + 1$  ou bien la forme finale, si  $l$  est le dernier niveau. Dans le cas où  $l = 1$ ,  $z(\sigma)$  prend en entrée le segment phonologique initial  $b_i$  indiqué dans l'entrée de  $i$ <sup>33 34</sup>.

Les cas de surabondance sont modélisés par la présence de plusieurs sous-schémas dans  $\mathbb{P}_i$ , puisque chacun d'eux ne produira qu'au plus une forme par PTM. Ce sont les sous-schémas qui explicitent les combinaisons licites de couples réalisationnels issus de différents niveaux. Ainsi, dans l'exemple de BALAYER à la figure 4.2, page 197, aucun sous-schéme ne permet la combinaison du radical BALAI avec la zone  $z_{v1,2}^{exp}$ , ce qui empêche de générer des formes erronées comme *\*balaions* ou *balaiez*. L'entrée lexicale associée à BALAYER est indiquée à la figure 4.24.

BALAYER	
I-PHON	balayer
I-CAT	<i>verbe</i>
MSF	{ <i>standard</i> }
S-STEM	( <i>vide</i> )
S-FORM	( <i>vide</i> )
I-PAT	( $z_{ay}^s$ , id), ( $z_{v1,1}^{exp}$ , id)
	( $z_{ay}^s$ , id), ( $z_{v1,2}^{exp}$ , id)
	( $z_{ai}^s$ , id), ( $z_{v1,2}^{exp}$ , id)

FIGURE 4.24 : Entrée lexicale pour le verbe BALAYER

Ssi au moins un PTM n'est associé à aucun couple réalisationnel dans au moins un sous-schéme du schème d'une entrée donnée, l'entrée en question présente un cas de DÉFECTIVITÉ FORMELLE, comme l'illustre l'entrée de TRAIRE indiquée plus haut à la figure 4.14, page 222.

33. Dans les implémentations de descriptions que nous présenterons dans le chapitre 5,  $b$  sera simplement la forme de citation d'une entrée lexicale donnée.

34. Nous reviendrons plus tard sur les cas de supplétion de formes ou de radicaux.



4.2.4.2 Zones réalisationnelles : illustration sur le verbe latin *aggredior*

Prenons l'exemple des verbes latins tels que décrits au chapitre 1. Selon notre analyse, un verbe latin se réalise selon des zones situées à trois niveaux réalisationnels distincts : le niveau radical, le niveau thématique et le niveau d'exponence. Plus précisément, un verbe déponent latin à thème mixte comme AGGREDIOR « *attaquer* » dont nous indiquons le paradigme au tableau 4.15<sup>35</sup> (page 240) sera associé aux zones réalisationnelles indiquées puis décrites ci-dessous. Chaque zone réalisationnelle  $y$  est indiquée avec sa règle de transfert associée qui en l'occurrence est à chaque fois la fonction identité, notée *id*.

- Niveau radical :  $(z_{ss,1}^s, id), (z_{ss,3}^s, id)$  ;
- Niveau thématique :  $(z_{c,2}^{th}, id), (z_{i,1}^{th}, id)$  ;
- Niveau d'exponence :  $(z_{D,1}^{exp}, id), (z_{D,3}^{exp}, id), (z_{C,3}^{exp}, id)$ .

Parmi les zones définissant la flexion d'AGGREDIOR, les zones de radicaux sont celles employées par les verbes construisant leurs deuxième et troisième radicaux en *-ss-*, c'est à dire appartenant à la classe de radicaux  $Z_{ss}^s$ . Plus précisément,  $z_{ss,1}^s$  correspond aux cases employant le RAD1 et  $z_{ss,3}^s$  aux cases employant le RAD3. AGGREDIOR n'ayant pas de formes du parfait syncrétiques qui sont les formes employant le RAD2, il n'est associé qu'aux zones  $z_{ss,1}^s$  et  $z_{ss,3}^s$  de la classe  $Z_{ss}^s$ , et pas à la zone  $z_{ss,2}^s$ . Les zones thématiques appartiennent pour  $z_{c,2}^{th}$  au thème des verbes athématiques et pour  $z_{i,1}^{th}$  au thème en *i*. Les zones flexionnelles quant à elles sont celles des formes imperfectives finies de la classe flexionnelle D ( $z_{D,1}^{exp}$ ) et celles des formes non finies de la classe flexionnelle C d'une part ( $z_{C,3}^{exp}$ ) et de la classe flexionnelle D ( $z_{D,3}^{exp}$ ) de l'autre. L'ensemble de ces zones est également illustré par les colorations respectives dans les tableaux 4.15 à 4.17 (pages 240 à 242), chaque couleur représentant une zone du niveau illustré par le tableau en question.

Dans les sous-schémas appartenant aux schémas de verbes comme AGGREDIOR « *attaquer* », ces couples réalisationnels sont arrangés par triplets indiquant comment les zones réalisationnelles des différents niveaux peuvent être combinées pour mener à la réalisation d'une forme d'un paradigme. Dans le cas de AGGREDIOR, les sous-schémas pertinents appartenant au schéma  $\mathbb{P}_{aggredior}$  de AGGREDIOR sont ceux indiqués dans l'entrée lexicale d'AGGREDIOR à la figure 4.25. Elle comporte par ailleurs la spécification de son appartenance à la catégorie flexionnelle verbale et de son inventaire standard de PTM.

35. Comme pour tous les verbes se fléchissant selon les classes flexionnelles B et C, les formes du parfait du verbe AGGREDIOR « *attaquer* » sont périphrastiques. Nous ne les indiquons donc pas dans le tableau 4.15 représentant le paradigme d'AGGREDIOR.

AGGREDIOR	
I-PHON	aggredior
I-CAT	<i>verbe</i>
MSF	{ <i>standard</i> }
S-STEM	( <i>vide</i> )
S-FORM	( <i>vide</i> )
I-PAT	( $z_{ss,1}^s$ , id), ( $z_{c,2}^{th}$ , id), ( $z_{D,1}^{exp}$ , id)
	( $z_{ss,1}^s$ , id), ( $z_{i,1}^{th}$ , id), ( $z_{D,1}^{exp}$ , id)
	( $z_{ss,1}^s$ , id), ( $z_{i,1}^{th}$ , id), ( $z_{C,3}^{exp}$ , id)
	( $z_{ss,3}^s$ , id), ( $z_{c,2}^{th}$ , id), ( $z_{C,3}^{exp}$ , id)
	( $z_{ss,1}^s$ , id), ( $z_{i,1}^{th}$ , id), ( $z_{D,3}^{exp}$ , id)
	( $z_{ss,3}^s$ , id), ( $z_{c,2}^{th}$ , id), ( $z_{D,3}^{exp}$ , id)

FIGURE 4.25 : Entrée lexicale pour le verbe latin AGGREDIOR

AGGREDIOR										
			CLASSE C				CLASSE D			
			IND		SBJV		IND		SBJV	
			SG	PL	SG	PL	SG	PL	SG	PL
IPFV	PRS	1					<i>aggredi-or</i>	<i>aggredi-mur</i>	<i>aggredi-ar</i>	<i>aggredi-amur</i>
		2					<i>aggrede-ris</i>	<i>aggredi-mini</i>	<i>aggredi-aris</i>	<i>aggredi-amini</i>
		3					<i>aggredi-tur</i>	<i>aggredi-ntur</i>	<i>aggredi-atur</i>	<i>aggredi-antur</i>
	PST	1					<i>aggredi-ebar</i>	<i>aggredi-ebamur</i>	<i>aggrede-rer</i>	<i>aggrede-remur</i>
		2					<i>aggredi-ebaris</i>	<i>aggredi-ebamini</i>	<i>aggrede-reris</i>	<i>aggrede-remini</i>
		3					<i>aggredi-ebatur</i>	<i>aggredi-ebantur</i>	<i>aggrede-retur</i>	<i>aggrede-rentur</i>
	FUT	1					<i>aggredi-ar</i>	<i>aggredi-emur</i>		
		2					<i>aggredi-eris</i>	<i>aggredi-emini</i>		
		3					<i>aggredi-etur</i>	<i>aggredi-entur</i>		
IMP	PRS	2				<i>aggrede-re</i>	<i>aggredi-mini</i>			
		2				<i>aggredi-tor</i>				
		3				<i>aggredi-tor</i>	<i>aggredi-ntor</i>			
INF	PRS						<i>aggredi</i>			
PTCP	PRS		<i>aggredi-ens, -entis</i>							
	PST							<i>aggress-us/-a/-um</i>		
GRV	FUT		<i>aggress-urus/-ura/-urum</i>							
			<i>aggredi-endi/-endo/-endum</i>					<i>aggredi-endus/-enda/-endum</i>		
GRD										
SUP			<i>aggress-us/-a/-um</i>							
							$z_{ss,1}^s$		$z_{ss,3}^s$	

TABLE 4.15 : Zones de radicaux dans le paradigme du verbe AGGREDIOR « *attaquer* »

AGGREDIOR										
			CLASSE C				CLASSE D			
			IND		SBJV		IND		SBJV	
			SG	PL	SG	PL	SG	PL	SG	PL
IPFV	PRS	1					<i>aggredi-or</i>	<i>aggredi-mur</i>	<i>aggredi-ar</i>	<i>aggredi-amur</i>
		2					<i>aggrede-ris</i>	<i>aggredi-mini</i>	<i>aggredi-aris</i>	<i>aggredi-amini</i>
		3					<i>aggredi-tur</i>	<i>aggredi-ntur</i>	<i>aggredi-atur</i>	<i>aggredi-antur</i>
	PST	1					<i>aggredi-tur</i>	<i>aggredi-ntur</i>	<i>aggredi-atur</i>	<i>aggredi-antur</i>
		2					<i>aggredi-ebar</i>	<i>aggredi-ebamur</i>	<i>aggrede-rer</i>	<i>aggrede-remur</i>
		3					<i>aggredi-ebaris</i>	<i>aggredi-ebamini</i>	<i>aggrede-reris</i>	<i>aggrede-remini</i>
	FUT	1					<i>aggredi-ebatur</i>	<i>aggredi-ebantur</i>	<i>aggrede-retur</i>	<i>aggrede-rentur</i>
		2					<i>aggredi-ar</i>	<i>aggredi-emur</i>		
		3					<i>aggredi-eris</i>	<i>aggredi-emini</i>		
IMP	PRS	2					<i>aggrede-re</i>	<i>aggredi-mini</i>		
		2					<i>aggredi-tor</i>			
		3					<i>aggredi-tor</i>	<i>aggredi-ntor</i>		
INF	PRS						<i>aggredi</i>			
PTCP	PRS					<i>aggredi-ens, -entis</i>				
	PST						<i>aggress-us/-a/-um</i>			
GRV	FUT					<i>aggress-urus/-ura/-urum</i>				
						<i>aggredi-endi/-endo/-endum</i>				
GRD							<i>aggredi-endus/-enda/-endum</i>			
SUP						<i>aggress-us/-a/-um</i>				

$z_{i,1}^h$

$z_{c,2}^h$

TABLE 4.16 : Zones thématiques dans le paradigme du verbe AGGREDIOR « *attaquer* »

AGGREDIOR										
			CLASSE C				CLASSE D			
			IND		SBJV		IND		SBJV	
			SG	PL	SG	PL	SG	PL	SG	PL
IPFV	PRS	1					<i>aggredi-or</i>	<i>aggredi-mur</i>	<i>aggredi-ar</i>	<i>aggredi-amur</i>
		2					<i>aggrede-ris</i>	<i>aggredi-mini</i>	<i>aggredi-aris</i>	<i>aggredi-amini</i>
		3					<i>aggredi-tur</i>	<i>aggredi-ntur</i>	<i>aggredi-atur</i>	<i>aggredi-antur</i>
	PST	1					<i>aggredi-ebar</i>	<i>aggredi-ebamur</i>	<i>aggrede-rer</i>	<i>aggrede-remur</i>
		2					<i>aggredi-ebaris</i>	<i>aggredi-ebamini</i>	<i>aggrede-reris</i>	<i>aggrede-remini</i>
		3					<i>aggredi-ebatur</i>	<i>aggredi-ebantur</i>	<i>aggrede-retur</i>	<i>aggrede-rentur</i>
	FUT	1					<i>aggredi-ar</i>	<i>aggredi-emur</i>		
		2					<i>aggredi-eris</i>	<i>aggredi-emini</i>		
		3					<i>aggredi-etur</i>	<i>aggredi-entur</i>		
IMP	PRS	2					<i>aggrede-re</i>	<i>aggredi-mini</i>		
		2					<i>aggredi-tor</i>			
		3					<i>aggredi-tor</i>	<i>aggredi-ntor</i>		
INF	PRS					<i>aggredi</i>				
PTCP	PRS		<i>aggredi-ens, -entis</i>							
	PST						<i>aggres-us/-a/-um</i>			
GRV	FUT		<i>aggres-urus/-ura/-urum</i>							
			<i>aggredi-endi/-endo/-endum</i>							
GRD							<i>aggredi-endus/-enda/-endum</i>			
SUP			<i>aggres-us/-a/-um</i>							

$\zeta_{D,1}^{exp}$

$\zeta_{D,1}^{exp}$

$\zeta_{D,3}^{exp}$

TABLE 4.17 : Zones flexionnelles dans le paradigme du verbe AGGREDIOR « *attaquer* »

#### 4.2.4.3 Fonctions de transfert et décalage morphosyntaxique

Une FONCTION DE TRANSFERT est une fonction qui prend en entrée un PTM et donne en sortie un autre PTM. Concrètement, les fonctions de transfert sont utilisées comme suit au cours de la réalisation des formes. Considérons un flexème  $i$  de schème  $\mathbb{P}_i$ . Notons  $\mathbb{P}$  l'un des sous-schémas de  $\mathbb{P}_i$ . La réalisation d'une forme au moyen de  $\mathbb{P}$  s'obtient par l'application successive des couples réalisationnels inclus dans  $\mathbb{P}$ . Soit  $p = (z, t)$  l'un d'entre eux. Dans le cas canonique,  $t$  est la fonction identité. Dans ce cas, la règle de réalisation associée par  $z$  à un PTM  $\sigma$  au niveau  $l$  est définie dès lors que  $\sigma$  appartient à l'ensemble de définition de  $z$ , et vaut simplement  $z(\sigma)$ . En revanche, si  $t$  n'est pas la fonction identité, la règle de réalisation associée par  $z$  à un PTM  $\sigma$  au niveau  $l$  n'est définie que si  $t(\sigma)$  appartient à l'ensemble de définition de  $z$ , et vaut alors  $z(t(\sigma))$ . Autrement dit, une fonction de transfert différente de l'identité permet de modéliser les situations non canoniques de DÉCALAGE MORPHOSYNTAXIQUE, où une zone est invoquée au cours de la réalisation flexionnelle avec un PTM distinct du PTM à réaliser.

Reprenons l'exemple des noms neutres du serbo-croate comme DETE « *enfant* » et BRAT « *frère* » illustrés par le tableau 4.11 page 216. Ils forment leur pluriel en empruntant les règles de réalisation du singulier des noms féminins illustrées tableau 4.10 page 216.

Les sous-schémas associés à DETE « *enfant* » et BRAT « *frère* » par leurs schémas respectifs sont donnés dans les entrées lexicales de DETE et BRAT aux figures 4.26 et 4.27 .

DETE	
I-PHON	dete
I-CAT	nom
MSF	{ <i>standard</i> }
S-STEM	( <i>vide</i> )
S-FORM	( <i>vide</i> )
I-PAT	( $Z_{\text{reg}}^s, \text{id}$ ), ( $z_{\text{N-ET, SG}}^{\text{exp}}, \text{id}$ )
	( $Z_{\text{reg}}^s, \text{id}$ ), ( $z_{\text{F-A, SG}}^{\text{exp}}, t_{\text{NB}}$ )

FIGURE 4.26 : Entrée lexicale pour le nom serbo-croate DETE « *enfant* »

Chacune des zones flexionnelles  $z_{\text{N-et,sg}}^{\text{exp}}$ ,  $z_{\text{M-a,sg}}^{\text{exp}}$  et  $z_{\text{F-a,sg}}^{\text{exp}}$  comprend un ensemble de règles de réalisation. Toutefois, les règles de réalisation de forme comprises dans ces zones sont toutes des règles de réalisation de PTM comprenant le {NOMBRE *sg*}. Or la zone  $z_{\text{F-a,sg}}^{\text{exp}}$

BRAT	
I-PHON	brat
I-CAT	<i>nom</i>
MSF	{ <i>standard</i> }
S-STEM	( <i>vide</i> )
S-FORM	( <i>vide</i> )
I-PAT	( $Z_{\text{reg}}^s, \text{id}$ ), ( $z_{M-A, SG}^{\text{exp}}, \text{id}$ )
	( $Z_{\text{reg}}^s, \text{id}$ ), ( $z_{F-A, SG}^{\text{exp}}, t_{\text{NB}}$ )

FIGURE 4.27 : Entrée lexicale pour le nom serbo-croate BRAT « frère »

a été sélectionnée pour réaliser les formes du pluriel de DETE « enfant » et BRAT « frère ». Les traits pluriels ne pouvant être réalisés par  $z_{F-a,sg}^{\text{exp}}$ , ce sont des PTM transformés qui sont donnés en entrées aux règles de réalisation comprises dans  $z_{F-a,sg}^{\text{exp}}$ .

Ces PTM transformés sont le résultat de l'application des fonctions de transfert associées à DETE et BRAT pour le niveau exposant. En l'occurrence il s'agit pour les deux flexèmes de la même fonction  $t_{\text{NB}}$ . Elle a pour effet de transformer chaque PTM en lui-même au nombre près : pour le nombre, elle transforme tout {NOMBRE *pl*} en {NOMBRE *sg*}. Ainsi, les PTM donnés en entrée aux règles de réalisation correspondent bien à ceux attendus par les règles de réalisation comprises dans les zones flexionnelles (associées aux entrées lexicales pour la réalisation du sous-paradigme du pluriel). Les traits effectivement réalisés (avant transformation) pour le flexème *i* étant des traits comportant le pluriel et les règles de réalisation étant des règles de réalisation du singulier, la réalisation échouerait en l'absence de la règle de transfert  $t_{\text{NB}}$  dans le couple réalisationnel concerné.

**Règle de sélection des couples réalisationnels, surabondance et défektivité** Étant donné un schème  $\mathbb{P}$ , définissons tout d'abord deux matrices dont les éléments contiennent respectivement seulement la zone et seulement la fonction de transfert rassemblés dans le couple réalisationnel  $\mathbb{P}_{j,k}$  :

$$\mathcal{Z}(\mathbb{P}) = \begin{pmatrix} z_{1,1} & z_{1,2} & z_{1,3} & \cdots & z_{1,m} \\ z_{2,1} & z_{2,2} & z_{2,3} & \cdots & z_{2,m} \\ z_{3,1} & z_{3,2} & z_{3,3} & \cdots & z_{3,m} \\ \vdots & & & & \\ z_{n,1} & z_{n,2} & z_{n,3} & \cdots & z_{n,m} \end{pmatrix}$$

$$\mathcal{T}(\mathbb{P}) = \begin{pmatrix} t_{1,1} & t_{1,2} & t_{1,3} & \cdots & t_{1,m} \\ t_{2,1} & t_{2,2} & t_{2,3} & \cdots & t_{2,m} \\ t_{3,1} & t_{3,2} & t_{3,3} & \cdots & t_{3,m} \\ \vdots & & & & \\ t_{n,1} & t_{n,2} & t_{n,3} & \cdots & t_{n,m} \end{pmatrix}$$

Soit maintenant  $i = (b_i, \Sigma_i, \mathcal{C}_i, \mathbb{P}_i, \mathcal{S}_i^{suppl}, F_i^{suppl})$  une entrée lexicale. Pour un niveau réalisationnel  $l$  de l'ensemble  $\{l_1, \dots, l_n\}$  des niveaux réalisationnels intervenant dans la flexion de  $i$ , nous définissons la RÈGLE DE SÉLECTION DE COUPLES RÉALISATIONNELS de  $i$  au niveau  $l$ , notée  $\mathbb{L}_l^i$ , comme étant l'ensemble des couples réalisationnels de niveau  $l$  dans  $\mathbb{P}_i$ . On peut noter que  $\mathbb{L}_l^i$  ne dépend que de  $\mathbb{P}_i$ . Si l'on note  $\text{proj}_l$  l'opération permettant ainsi d'extraire l'intégralité des couples réalisationnels utilisés par un schème au niveau  $l$ , on a alors  $\mathbb{L}_l^i = \text{proj}_l(\mathbb{P}_i)$ .

Nous définissons enfin l'ensemble  $\mathbb{L}_{l|Z}^i$  des zones réalisationnelles intervenant dans la flexion d'un flexème  $i$  comme suit :  $\mathbb{L}_{l|Z}^i = \{z | \exists t | (z, t) \in \mathbb{L}_l^i\}$ . En étendant de façon intuitive l'opérateur  $\text{proj}_l$ , on peut également écrire que  $\mathbb{L}_{l|Z}^i = \text{proj}_l(\mathcal{Z}(\mathbb{P}_i))$ .

Soit  $\sigma$  un PTM appartenant à  $\Sigma_i$ ,  $l$  un niveau réalisationnel et  $\mathbb{L}_l^i$  la règle de sélection de couples réalisationnels pour le niveau  $l$ . Dans le cas canonique, une et une seule forme réalise le PTM  $\sigma$ , ce qui veut dire que pour tout niveau  $l$ , la règle de sélection de couples réalisationnels  $\mathbb{L}_l^i$  associée à  $\sigma$  un ensemble de couples réalisationnels réduit à un singleton. À l'inverse, si, pour l'un des  $\sigma \in \Sigma_i$  et l'un des niveaux  $l$ ,  $\mathbb{L}_l^i(\sigma)$  est un ensemble contenant plus d'un élément, la case correspondant au PTM  $\sigma$  est SURABONDANTE. À l'inverse, les cas où  $\mathbb{L}_l^i(\sigma)$  est l'ensemble vide conduisent à une DÉFECTIVITÉ formelle en  $\sigma$  au niveau  $l$ .

Dans l'exemple du verbe BALAYER indiqué page 197,  $\mathbb{L}_{l^{exp}}^{\text{BAL}}$  sélectionne ainsi les zones  $z_{v1,1}^{exp}$  et  $z_{v1,2}^{exp}$  au niveau exposant  $l^{exp}$ .



### 4.2.5 Construction des formes et phénomènes de supplétion

Comme nous l'avons rappelé ci-dessus, un radical  $s$  dans  $\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}$  est un couple de la forme  $(z^s, r^s)$ , où  $z^s$  est une zone de radicaux de niveau 1 et  $r^s$  la réalisation (chaîne de phonèmes ou de graphèmes) finale de la suite des règles de réalisation de niveaux radical ou thématiques. Une forme  $f$  est un couple de la forme  $(\sigma, r^f)$  où  $\sigma$  est un PTM et  $r^f$  la réalisation finale de l'ensemble des règles de réalisation appliquées à un flexème donné pour un PTM particulier.

Soit une entrée lexicale  $i = (b_i, \Sigma_i, \mathcal{C}_i, \mathbb{P}_i, S_i^{suppl}, F_i^{suppl})$  et  $\sigma \in \Sigma_i$  un PTM qui correspond donc à une case non vide du paradigme de  $i$ . Dans le cas canonique,  $S_i^{suppl}$  et  $F_i^{suppl}$  sont vides, et une forme  $f$  de la case de  $i$  associée au PTM  $\sigma \in \Sigma_i$  est un couple formé d'une part de  $\sigma$  et d'autre part du résultat de l'application en  $\sigma$  des couples réalisationnels associés à  $\sigma$  par l'un des sous-schémas de  $\mathbb{P}_i$  à l'ensemble des niveaux, de la façon suivante.

Soit  $l$  un niveau réalisationnel appartenant à l'ensemble  $\{l_1, \dots, l_n\}$  des niveaux intervenant dans la flexion de  $i$ . La règle de sélection de zones de niveau  $l$ ,  $\mathbb{L}_l^i$ , associée à  $\sigma$  un ensemble de couples réalisationnels. Parmi eux, considérons l'un d'entre eux, noté  $(z, t)$ . Appliquer  $z$  à  $\sigma$  permet alors d'obtenir la règle de réalisation  $R = z(t(\sigma))$  sélectionnée par le couple réalisationnel retenu. La règle réalisationnelle  $R$  prend en entrée une réalisation de niveau  $l - 1$  si  $l > 1$ , et la séquence phonologique de base  $b_i$  de  $i$  sinon. Si le flexème  $i$  est surabondant à deux niveaux différents  $l_1$  et  $l_2$  pour un même PTM  $\sigma$ , rien ne garantit que tous les couples réalisationnels de  $\mathbb{L}_{l_1}^i(\sigma)$  et tous les couples réalisationnels de  $\mathbb{L}_{l_2}^i(\sigma)$  puissent participer ensemble à la construction de formes de  $i$  réalisant  $\sigma$ . Ce sera l'objet des sous-schémas que d'explicitier les séquences licites de règles de réalisation.

Si  $i$  présente des cas de supplétion de radical au sens strict<sup>36</sup>, elle comporte un ENSEMBLE DE RADICAUX SUPPLÉTIFS  $S_i^{suppl}$  non vide. De même, si  $i$  présente des cas de supplétion de forme, elle comporte un ENSEMBLE DE FORMES SUPPLÉTIVES  $F_i^{suppl}$  non vide. Nous avons vu des exemples de telles listes pour les verbes ALLER et ÊTRE respectivement aux figures 4.5 et 4.7 indiquées plus haut, pages 204 et 203.

Si  $F_i^{suppl}$  n'est pas vide, soit  $\sigma$  un PTM tel qu'il existe au moins un élément  $f$  de la forme  $(\sigma, r^f)$  dans  $F_i^{suppl}$ . Alors les seules formes présentes dans la case du paradigme de  $i$  associée au PTM  $\sigma$  sont les formes  $f = (\sigma, r^f)$  telles que  $(\sigma, r^f) \in F_i^{suppl}$ . Pour les PTM tels qu'il n'existe pas d'élément de la forme  $(\sigma, r^f)$  dans  $F_i^{suppl}$ , la construction des formes pour la case correspondant à  $\sigma$  se fait comme dans le cas canonique, sauf dans le cas d'une supplétion de radical, cas où l'on procède comme suit.

36. Les cas où l'allomorphie de radicaux n'est pas obtenue par les règles de formation de radicaux régulières comprises dans la grammaire.

Si  $S_i^{suppl}$  n'est pas vide, soit  $s = (z^s, r^s)$  l'un de ses élément et  $\pi_{z^s}$  l'espace partitionnant qui constitue l'espace de définition de la zone radicale  $z^s$ . Soit enfin un sous-schème  $\mathbb{p}$  inclus dans  $\mathbb{P}_i$  qui associe au niveau radical 1 un couple réalisationnel dont la zone radicale est précisément  $z^s$ . Alors la réalisation d'une forme pour un PTM  $\sigma \in \pi_{z^s}$  au moyen du sous-schème  $\mathbb{p}$  ne fera pas usage du couple réalisationnel correspondant : les règles de réalisation de type exponence (appartenant aux niveaux d'exponence) s'appliqueront directement à  $s$ .

#### 4.2.6 Les classes flexionnelles et les classes de radicaux

De même que pour les catégories flexionnelles, les propriétés des paradigmes observables permettent d'identifier parmi les zones réalisationnelles existantes celles qui, niveau par niveau, tendent à être associées conjointement aux entrées lexicales d'une langue donnée. Ces zones, comportant habituellement les règles de réalisation d'un même paradigme, seront appelées CLASSES RÉALISATIONNELLES dans  $\mathcal{PARSL}$ . Les classes flexionnelles (ou d'exponence) comme les classes de radicaux et les classes thématiques (ou thèmes) ne constituent ainsi pas des primitives de la grammaire dans  $\mathcal{PARSL}$ , mais sont le reflet d'un comportement majoritaire des flexèmes de la langue étudiée.

L'HÉTÉROCLISE correspond alors à l'association d'un flexème donné avec des zones n'appartenant pas, par défaut, à des classes réalisationnelles établies.

### 4.3 Quelques caractéristiques saillantes de $\mathcal{PARSL}$ comparées à PFM et *Network Morphology*

**La notion de zone réalisationnelle** Une des caractéristiques les plus saillantes de  $\mathcal{PARSL}$  est sa généralisation de la notion de zone de radicaux, issue de la tradition des espaces partitionnants de Pirrelli & Battista (2000) et des espaces thématiques de Bonami & Boyé (2003); Bonami et al. (2008); Bonami & Boyé (2010a), à l'ensemble du mécanisme de la flexion. Cette propriété démarque notamment  $\mathcal{PARSL}$  du fonctionnement de la correspondance paradigmatique proposée par Stump (2006) pour PFM. En PFM, c'est le radical d'un lexème, et non pas le lexème lui-même, qui est associé à une classe flexionnelle. Cela permet également d'associer un lexème donné avec une sous-partie d'une classe flexionnelle, dans la mesure où seules les règles réalisationnelles compatibles avec les cases pour lesquelles un radical donné a été sélectionné pourront être appliquées au radical en question pour la construction des formes. Par exemple, dans une analyse PFM, un lexème verbal du français en *-ayer* serait associé à deux radicaux, l'un en *-ay*, l'autre en *-ai*.

Chacun de ces radicaux serait associé à la classe flexionnelle des verbes du premier groupe. Pour chacun de ces radicaux, seules les règles compatibles avec les cases pour lesquelles le radical en *-ai*, respectivement en *-ay*, seraient appliquées au radical en question. En fin de compte, l'intersection entre les traits exprimés par les règles et les cases compatibles suite à la sélection de radicaux correspond justement aux espaces partitionnants définissant les zones d'exponence que nous avons posées dans notre analyse  $\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}$ . Mais ces zones flexionnelles n'ont pas en PFM de statut théorique en tant que telles.

En *Network Morphology*, les lexèmes sont directement associés à des classes flexionnelles, comme c'est le cas dans  $\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}$ .  $\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}$  partage également avec *Network Morphology* le fait qu'une entrée puisse stipuler plusieurs radicaux différents qui seront combinés avec les règles des réalisations de forme selon leur compatibilité. En revanche, *Network Morphology* ne comporte pas la notion de zone flexionnelle. Les lexèmes sont associés à des classes complètes. On pourrait reconstruire la notion de zone flexionnelle en partitionnant la classe flexionnelle en deux sous-classes. Chaque entrée lexicale ne serait alors pas associée à une classe flexionnelle mais à deux sous-classes. La solution changerait cependant la définition de la notion de classe flexionnelle qui pourtant constitue une primitive de *Network Morphology*.

**Généralisation de la notion de classe** Une autre propriété particulière de  $\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}$  est qu'il généralise la notion de classe flexionnelle aux notions de classe thématique et de classe de radicaux.

Comme mentionné plus haut, ces notions généralisent les notions d'espace thématique et d'espace partitionnant. Mais ces deux approches, fondées essentiellement sur l'étude des langues romanes, n'envisagent pas une alternance radicale régulière indépendante d'une alternance thématique au sens de  $\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}$ . Lorsque cela apparaît nécessaire pour l'analyse d'une langue donnée,  $\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}$  permet ainsi de rendre compte séparément des alternances radicales proprement dites<sup>37</sup> et des alternances thématiques. La généralisation de la notion de classe à tous les niveaux réalisationnels a également l'avantage de rendre possible l'identification de phénomènes d'hétéroclise à l'ensemble de ces niveaux.

Les classes de radicaux, de thème et d'exponence tels qu'elles sont définies dans  $\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}$  seraient intégrables à peu de frais à *Network Morphology* par une multiplication des hiérarchies morphologiques de sorte à intégrer une hiérarchie de classes séparées pour les radicaux d'une part et les thèmes de l'autre. Ces hiérarchies parallèles hériteraient alors les unes des autres par des liens d'héritage orthogonaux. Une telle approche correspondrait à

37. Voir les données du persan à la figure 4.4 page 200.

une généralisation des hiérarchies parallèles exprimables pour les exposants et notamment employées pour la description de la flexion des noms russes (Brown et al., 1996 ; Brown & Hippisley, 2012). Différentes hiérarchies d'exponence y sont employées pour représenter l'exponence exprimée par suffixation d'une part et par des propriétés accentuelles de l'autre. Leurs interactions sont captées par l'intermédiaire d'héritages orthogonaux entre nœuds de ces hiérarchies parallèles. Le même principe que celui appliqué aux exposants dans la description de la flexion nominale du russe par Brown et al. (1996) et Brown & Hippisley (2012) serait généralisé à tous les niveaux, c'est-à-dire, dans le cas de l'interaction entre niveaux de radicaux de thèmes et d'exposants, à tous les éléments intervenant dans la réalisation des formes d'un paradigme.

En PFM, en revanche, la bi-partition entre paradigme de contenu et paradigme de forme ne semble pas compatible avec la notion de classe de radicaux. Le fait que les radicaux soient individuellement associés à des classes flexionnelles et possèdent chacun leur paradigme propre semble en effet incompatible avec une relation paradigmatique entre les radicaux eux-mêmes.

**Représentation immédiate du décalage morphosyntaxique** PARSLI permet enfin de rendre compte par les règles de transfert indiquées dans la structure d'une entrée lexicale particulière des phénomènes de décalages morphosyntaxiques observables dans son paradigme. En *Network Morphology*, le décalage est pris en compte par un mécanisme analogue : les règles de renvoi. Elles stipulent pour des entrées comme celles du serbo-croate que pour réaliser le pluriel il faut suivre les réalisations stipulées pour le singulier des noms féminin en *-a*. Toutefois, ces règles de renvoi sont un mécanisme générique en *Network Morphology*. Elles n'y figurent pas en tant que primitives destinées à encoder ce phénomène non canonique particulier. En PFM, le décalage est pris en compte par une différence entre le paradigme de contenu et le paradigme de forme, réalisé par une règle de correspondance paradigmatique non standard<sup>38</sup>.

## 4.4 Quantifier la différence : les mesures de canonicité dans PARSLI

Comme nous l'avons montré dans les sections 4.1 et 4.2, PARSLI modélise explicitement les phénomènes de flexion non canonique. Nous pouvons de surcroît définir une série

---

38. Voir la description de ce mécanisme à la page 133 du chapitre 2.

de mesures de canonicité pour chacun des phénomènes non canoniques. Ces mesures permettent de rendre compte de l'étendue de la non-canonicité d'un paradigme ainsi que, appliquée à un système flexionnel d'une langue donnée, de l'intégralité de ce système. Étant donné que chacune des primitives a été spécifiquement conçue pour rendre compte des propriétés canoniques des paradigmes modélisés, les mesures de canonicité que nous présentons dans cette section sont pour la plupart très simples et constituent de simples comptages des déviations par rapport au canon. Cette simplicité des mesures découle ainsi directement de l'adéquation entre les phénomènes modélisés et le modèle formel lui-même.

Dans cette section, nous introduisons et nous illustrons les mesures de canonicité que nous avons conçues. L'application de ces mesures à des descriptions morphologiques quasi-exhaustives d'un système flexionnel particulier fera l'objet du chapitre 7. Nous y verrons comment ces mesures permettent de caractériser le degré de canonicité d'un système flexionnel dans son ensemble.

Les notations employées dans cette partie sont celles indiquées à la figure 4.18 page 229.

Les mesures que nous proposons s'appuient sur la définition d'une entrée lexicale illustrée par la figure 4.19 page 231.

#### 4.4.1 Mesures du remplissage des cases du paradigme

Un premier type de mesures concerne le remplissage plus ou moins canonique des cases du paradigme d'un flexème donné. Elles s'appuient sur la définition du schème d'un flexème et des zones de sélection de règles de réalisation associées à un flexème donné pour un niveau donné. Parmi les phénomènes concernés figurent la DÉFECTIVITÉ FORMELLE, la SURABONDANCE, l'ALLOMORPHIE RÉGULIÈRE et l'HÉTÉROCLISE. Les indices mesurant chacun de ces phénomènes sont à chaque fois définis pour un paradigme particulier et pour l'ensemble d'un système flexionnel.

##### 4.4.1.1 Indice de défektivité

**Défektivité formelle d'un paradigme** Intuitivement, un paradigme est défektivif s'il comporte des cases que l'entrée lexicale du flexème en question ne peut pas fléchir. On appellera ces cases des CASES DÉFECTIVES.

Plus formellement, le paradigme d'un flexème  $i$  est défektivif si et seulement s'il existe au moins un PTM  $\sigma \in \Sigma_{\mathcal{C}_i}$  tel que  $\sigma \in \Sigma_i$ , et un niveau  $l$  appartenant à l'ensemble  $\{l_{\mathcal{C}_i,1}, \dots, l_{\mathcal{C}_i,n}\}$  des niveaux réalisationnels intervenant canoniquement dans la flexion des flexèmes de catégorie flexionnelle  $\mathcal{C}_i$  pour lequel l'ensemble  $\mathbb{L}_l^i(\sigma)$  de couples

réalisationnels est vide.

$$\exists \sigma \in \Sigma_{\mathcal{C}_i} \cap \Sigma_i \mid \exists l \in \{l_{\mathcal{C}_i,1}, \dots, l_{\mathcal{C}_i,n}\} \mid \mathbb{L}_l^i(\sigma) = \emptyset \quad (4.1)$$

Reprenons l'exemple de TRAIRE que nous avons donné à la figure 4.14 page 222. Un exemple de PTM qui dans aucun des sous-schémas n'est associé à un couple réalisationnel de type radical est celui de la première personne du singulier du passé simple de l'indicatif. Nous rappelons l'entrée lexicale correspondante à la figure 4.28.

TRAIRE							
I-PHON	traire						
I-CAT	<i>verbe</i>						
MSF	{ <i>standard</i> }						
S-STEM	( <i>vide</i> )						
S-FORM	( <i>vide</i> )						
I-PAT	<table style="border-collapse: collapse; margin-left: 20px;"> <tr> <td style="padding-right: 10px;">(z<sub>ai</sub><sup>s</sup>, id),</td> <td>(Z<sub>TRAIRE</sub><sup>exp</sup>, id)</td> </tr> <tr> <td style="padding-right: 10px;">(z<sub>ay</sub><sup>s</sup>, id),</td> <td>(Z<sub>TRAIRE</sub><sup>exp</sup>, id)</td> </tr> <tr> <td style="padding-right: 10px;">—</td> <td>(Z<sub>TRAIRE</sub><sup>exp</sup>, id)</td> </tr> </table>	(z <sub>ai</sub> <sup>s</sup> , id),	(Z <sub>TRAIRE</sub> <sup>exp</sup> , id)	(z <sub>ay</sub> <sup>s</sup> , id),	(Z <sub>TRAIRE</sub> <sup>exp</sup> , id)	—	(Z <sub>TRAIRE</sub> <sup>exp</sup> , id)
(z <sub>ai</sub> <sup>s</sup> , id),	(Z <sub>TRAIRE</sub> <sup>exp</sup> , id)						
(z <sub>ay</sub> <sup>s</sup> , id),	(Z <sub>TRAIRE</sub> <sup>exp</sup> , id)						
—	(Z <sub>TRAIRE</sub> <sup>exp</sup> , id)						

FIGURE 4.28 : Entrée lexicale déficiente : le verbe TRAIRE

**Indice de défektivité formelle d'un paradigme** L'indice de défektivité formelle d'un paradigme se calcule alors comme le nombre de cases déficientes divisé par le nombre de cases du paradigme.

Pour un flexème  $i = (b_i, \Sigma_i, \mathcal{C}_i, \mathbb{P}_i, S_i^{suppl}, F_i^{suppl})$  donné, l'indice de défektivité formelle  $\mathcal{F}^-(i)$  de son paradigme se calcule de la façon suivante :

$$\mathcal{F}^-(i) = \frac{|\{\sigma \in \Sigma_{\mathcal{C}_i} \cap \Sigma_i \mid \exists l \in \{l_1, \dots, l_n\} \mid \mathbb{L}_l^i(\sigma) = \emptyset\}|}{|\Sigma_{\mathcal{C}_i}|} \quad (4.2)$$

L'indice de défektivité formelle du paradigme de TRAIRE est ainsi égal au nombre de PTM appartenant à l'ensemble  $\Sigma_{\text{TRAIRE}}$  non associés à un couple réalisationnel par la règle de sélection de couples réalisationnels  $\mathbb{L}_l^{\text{TRAIRE}}(\sigma)$  associée à TRAIRE divisé par le nombre

de PTM canoniquement exprimés par un flexème verbal du français. En l'occurrence, selon le Bescherelle (Arrivé, 1997), les PTM concernés sont les six PTM du passé simple de l'indicatif et les six PTM de l'imparfait du subjonctif. Or le nombre de PTM exprimés par une forme synthétique dans les paradigmes canoniques des flexèmes verbaux du français est de 48<sup>39</sup>. L'indice de défektivité formelle du paradigme de TRAIRE vaut donc :

$$\begin{aligned} \mathcal{F}^-(\text{TRAIRE}) &= \frac{|\{\sigma \in \Sigma_{\text{VERBE}} \cap \Sigma_{\text{TRAIRE}} \mid \exists l \in \{l_{\text{VERBE},1}, \dots, l_{\text{VERBE},n}\} \mid \mathbb{L}_l^{\text{TRAIRE}}(\sigma) = \emptyset\}|}{|\Sigma_{\text{VERBE}}|} \\ &= \frac{6+6}{48} = \frac{12}{48} \\ &= 25\% \end{aligned} \tag{4.3}$$

**Indice de défektivité formelle d'un système flexionnel** Nous définissons l'indice de défektivité formelle d'un système flexionnel  $\mu_{\text{infl}}$  comme la moyenne des indices de défektivité associés à chacun de ses paradigmes sur le lexique flexionnel.

L'indice de défektivité formelle  $\mathcal{F}^-$  de  $\mu_{\text{infl}}$  est donc obtenu comme suit :

$$\mathcal{F}^-(\mu_{\text{infl}}) = \frac{\sum_{i \in \mathcal{I}} \mathcal{F}^-(i)}{|\mathcal{I}|} \tag{4.4}$$

#### 4.4.1.2 Indice de surabondance

**Surabondance d'un paradigme** Le paradigme d'un flexème  $i$  sera considéré comme étant surabondant si et seulement s'il existe au moins un PTM appartenant à l'ensemble  $\Sigma_i$  des PTM pour lesquels  $i$  est défini, qui est associé à plus d'un couple réalisationnel par la règle de sélection de couples réalisationnels d'au moins un niveau  $l$  appartenant à l'ensemble  $\{l_{\mathcal{C}_i,1}, \dots, l_{\mathcal{C}_i,n}\}$  des niveaux réalisationnels canoniques pour sa catégorie  $\mathcal{C}_i$ . Les cases correspondantes seront dites CASES SURABONDANTES.

Plus précisément, le paradigme d'un flexème  $i$  est surabondant si et seulement si :

$$\exists \sigma \in \Sigma_{\mathcal{C}_i} \cap \Sigma_i \mid \exists l \in \{l_{\mathcal{C}_i,1}, \dots, l_{\mathcal{C}_i,n}\} \mid |\mathbb{L}_l^i(\sigma)| > 1 \tag{4.5}$$

39. Le chiffre que nous indiquons ici correspond au nombre de formes indiquées par le Bescherelle (Arrivé, 1997). Il n'indique à chaque fois qu'une seule forme pour le participe passé. Compter les quatre formes du participe passé porterait le nombre de formes dans les paradigmes verbaux canoniques du français à 51.

Dans le cas de ASSEOIR que nous avons mentionné au chapitre 3 et dont nous rappelons un extrait du paradigme au tableau 4.18, la surabondance concerne 34 PTM sur les 48 canoniquement réalisés par les verbes du français (Arrivé, 1997). Nous donnons l'entrée lexicale simplifiée du verbe ASSEOIR à la figure 4.29. Dans les quatre premiers sous-schémas, le premier et le deuxième sous-schéma d'une part et le troisième et le quatrième sous-schéma de l'autre servent à chaque fois à la réalisation des mêmes PTM parmi les 34 PTM doublement réalisés.

IND.PRS	SG	PL
1	<i>assois</i>	<i>asseyons</i>
	<i>assieds</i>	<i>asseyons</i>
2	<i>assois</i>	<i>asseyez</i>
	<i>assieds</i>	<i>asseyez</i>
3	<i>assoit</i>	<i>assoient</i>
	<i>assied</i>	<i>asseyent</i>

TABLE 4.18 : Surabondance en français pour le verbe ASSEOIR

ASSEOIR		
I-PHON	asseoir	
I-CAT	<i>verbe</i>	
MSF	{ <i>standard</i> }	
S-STEM	( <i>vide</i> )	
S-FORM	( <i>vide</i> )	
I-PAT	$(z_{ied}^s, id)$ ,	$(Z_{ASSEOIR}^{exp}, id)$
	$(z_{oi}^s, id)$ ,	$(Z_{ASSEOIR}^{exp}, id)$
	$(z_{oy}^s, id)$ ,	$(Z_{ASSEOIR}^{exp}, id)$
	$(z_{ey}^s, id)$ ,	$(Z_{ASSEOIR}^{exp}, id)$
	$(z_i^s, id)$	$(Z_{ASSEOIR}^{exp}, id)$

FIGURE 4.29 : Entrée lexicale surabondante : le verbe ASSEOIR

**Indice de surabondance d'une case à un niveau donné** L'indice de surabondance à un niveau  $l \in \{l_{c_i,1}, \dots, l_{c_i,n}\}$  donné d'une case  $C_\sigma$  définie pour un PTM  $\sigma$  appartenant à  $\Sigma_i$



dans le paradigme d'un flexème  $i$  se calcule alors comme le nombre de zones surnuméraires associées à  $\sigma$  par la règle de sélection couples réalisationnels  $\mathbb{L}_l^i \in \{\mathbb{L}_{l_1}^i, \dots, \mathbb{L}_{l_n}^i\}$ . Ce nombre est le nombre de zones associées au PTM  $\sigma$  en question moins 1.

Pour un PTM  $\sigma \in \Sigma_i$  et un niveau  $l \in \{l_{e_i,1}, \dots, l_{e_i,n}\}$ , on définit l'indice de surabondance  $\mathcal{F}_l^+(i, \sigma)$  d'un flexème  $i$  dans une case  $C_\sigma$  de son paradigme comme suit :

$$\mathcal{F}_l^+(i, \sigma) = |\mathbb{L}_l^i(\sigma)| - 1 \quad (4.6)$$

Telle que nous avons représenté l'entrée lexicale du verbe ASSEOIR à la figure 4.29, la surabondance de formes dans le paradigme provient d'une surabondance au niveau réalisationnel radical. Pour le 1SG.PRS.IND, l'entrée lexicale d'ASSEOIR comporte deux sous-schémes  $(z_{ied}^s, \text{id})$ ,  $(Z_{ASSEOIR}^{exp}, \text{id})$  et  $(z_{oi}^s, \text{id})$ ,  $(Z_{ASSEOIR}^{exp}, \text{id})$  de sorte que 1SG.PRS.IND est associé à la fois au couple réalisationnel  $(z_{ied}^s, \text{id})$  et au couple réalisationnel  $(z_{oi}^s, \text{id})$  par la règle de sélection de couples réalisationnels  $\mathbb{L}_{l_s}^{ASSEOIR}$  du niveau radical.

L'indice de surabondance de la case du 1SG.PRS.IND du paradigme d'ASSEOIR au niveau radical vaut ainsi :

$$\begin{aligned} \mathcal{F}_{l_s}^+(ASSEOIR, 1SG.PRS.IND) &= |\mathbb{L}_{l_s}^{ASSEOIR}(1SG.PRS.IND)| - 1 \\ &= 2 - 1 \\ &= 1 \end{aligned} \quad (4.7)$$

**Indice de surabondance globale d'une case dans un paradigme** L'indice de surabondance globale d'une case dans un paradigme d'un flexème  $i = (b_i, \Sigma_i, \mathcal{C}_i, \mathbb{P}_i, S_i^{suppl}, F_i^{suppl})$  est nul si cette case est vide. Sinon, il est égal au nombre de formes qu'elle contient moins 1. Le nombre de formes contenues dans une case est noté  $|C_\sigma|$ .

Pour un PTM  $\sigma \in \Sigma_i$ , l'indice de surabondance globale  $\mathcal{F}_C^+(i, \sigma)$  de la case  $C_\sigma$  qui lui correspond est calculé comme suit :

$$\mathcal{F}_C^+(i, \sigma) = \max(0, |C_\sigma| - 1) \quad (4.8)$$

Pour le verbe ASSEOIR, le 1SG.PRS.IND est associé à deux zones réalisationnelles. La case du 1SG.PRS.IND comporte deux formes *assieds* et *assois*. Ainsi l'indice de surabondance globale du 1SG.PRS.IND dans le paradigme du verbe ASSEOIR vaut :

$$\begin{aligned}
 \mathcal{F}_C^+(\text{ASSEOIR}, \text{1SG.PRS.IND}) &= \max(0, |C_{\text{1SG.PRS.IND}}| - 1) \\
 &= \max(0, 2 - 1) = \max(0, 1) \\
 &= 1
 \end{aligned} \tag{4.9}$$

Notons que dans le cas d'ASSEOIR, la surabondance provient exclusivement d'une surabondance au niveau radical. Le nombre de formes par case est donc égal au nombre de radicaux dans cette case. L'indice de surabondance globale du 1SG.PRS.IND est ainsi égal à l'indice de surabondance de 1SG.PRS.IND à l'unique niveau radical.

**Indice de dispersion de la surabondance d'une case dans un paradigme à travers les niveaux** L'indice de dispersion de la surabondance d'une case  $C_\sigma$  dans un paradigme à travers les niveaux permet de mesurer le degré de généralité du phénomène de surabondance de cette case à travers les niveaux réalisationnels. Il correspond au nombre de niveaux auxquels  $C_\sigma$  est surabondante. Il est ainsi défini comme le nombre de niveaux surabondants en cette case divisé par le nombre de niveaux.

Pour un PTM  $\sigma \in \Sigma_i$ , l'indice de dispersion  $\mathcal{F}_{C,disp}^+(i, \sigma)$  de la surabondance d'une case  $C_\sigma$  dans un paradigme à travers les niveaux est calculé comme suit :

$$\mathcal{F}_{C,disp}^+(i, \sigma) = \frac{|\{l \in \{l_{\mathcal{C}_i,1}, \dots, l_{\mathcal{C}_i,n}\} \mid \mathcal{F}_l^+(i, \sigma) \neq 0\}|}{|\{l_{i,1}, \dots, l_{i,n}\}|} \tag{4.10}$$

Pour ASSEOIR l'indice de dispersion de la case du 1SG.PRS.IND à travers les niveaux vaut ainsi :

$$\begin{aligned}
 \mathcal{F}_{C,disp}^+(\text{ASSEOIR}, \text{1SG.PRS.IND}) &= \frac{|\{l \in \{l_{\text{VERBE},1}, \dots, l_{\text{VERBE},n}\} \mid \mathcal{F}_l^+(\text{ASSEOIR}, \text{1SG.PRS.IND}) \neq 0\}|}{|\{l_{\text{ASSEOIR},1}, \dots, l_{\text{ASSEOIR},n}\}|} \\
 &= \frac{1}{2} \\
 &= 50\%
 \end{aligned} \tag{4.11}$$

**Indice de dispersion de la surabondance dans un paradigme** L'indice de dispersion de la surabondance dans un paradigme permet de mesurer le degré de généralité du phénomène dans le paradigme. Il est défini comme le nombre de cases surabondantes (c'est-

à-dire dont l'indice de surabondance globale n'est pas nul) divisé par la taille du paradigme, c'est-à-dire son nombre de cases.

L'indice de dispersion  $\mathcal{F}_{P,disp}^+(i)$  de la surabondance dans un paradigme  $P_i$  est alors calculé comme suit :

$$\mathcal{F}_{P,disp}^+(i) = \frac{|\{\sigma \in \Sigma_i | \mathcal{F}_C^+(i, \sigma) \neq 0\}|}{|\Sigma_i|} \quad (4.12)$$

Pour le verbe ASSEOIR, le nombre de cases sur abondante dans son paradigme est de 34. L'indice de dispersion  $\mathcal{F}_{P,disp}^+(\text{ASSEOIR})$  de la surabondance dans son paradigme  $P_{\text{ASSEOIR}}$  vaut ainsi :

$$\begin{aligned} \mathcal{F}_{P,disp}^+(\text{ASSEOIR}) &= \frac{|\{\sigma \in \Sigma_{\text{ASSEOIR}} | \mathcal{F}_C^+(\text{ASSEOIR}, \sigma) \neq 0\}|}{|\Sigma_{\text{ASSEOIR}}|} \\ &= \frac{34}{48} \\ &\approx 71\% \end{aligned} \quad (4.13)$$

**Indice de la surabondance moyenne d'un paradigme** Si le nombre de cases surabondantes est nul, l'indice de surabondance moyenne est défini comme étant nul. Sinon, l'indice de surabondance moyenne d'un paradigme équivaut à la somme des indices de surabondance globale des cases du paradigme divisée par le nombre de cases du paradigme.

Formellement :

$$\mathcal{F}_{P,av}^+(i) = \frac{\sum_{\sigma \in \Sigma_i} \mathcal{F}_C^+(i, \sigma)}{|\Sigma_i|} \quad (4.14)$$

Dans le cas d'ASSEOIR, l'indice de la surabondance moyenne de son paradigme revient à :

$$\begin{aligned}
 \mathcal{F}_{P,av}^+(\text{ASSEOIR}) &= \frac{\sum_{\sigma \in \Sigma_{\text{ASSEOIR}}} \mathcal{F}_C^+(\text{ASSEOIR}, \sigma)}{|\Sigma_{\text{ASSEOIR}}|} \\
 &= \frac{34 \times 1}{48} \\
 &\approx 71\%
 \end{aligned}
 \tag{4.15}$$

**Indice de la surabondance moyenne dans le système flexionnel** Si aucun flexème n'est surabondant, l'indice de la surabondance moyenne d'un système flexionnel est défini comme étant nul. Sinon, l'indice de la surabondance moyenne d'un système flexionnel est la somme des indices de surabondance moyenne des paradigmes des lexèmes de son lexique flexionnel divisée par la taille du lexique flexionnel.

L'indice de la surabondance moyenne  $\mathcal{F}_{av}^+(\mu_{\text{infl}})$  du système flexionnel  $\mu_{\text{infl}}$  est calculé comme suit :

$$\mathcal{F}_{av}^+(\mu_{\text{infl}}) = \frac{\sum_{i \in \mathcal{L}} \mathcal{F}_{P,av}^+(i)}{|\mathcal{L}|}
 \tag{4.16}$$

#### 4.4.1.3 Indice d'allomorphie

**Allomorphie radicale d'un flexème** Le degré d'allomorphie radicale d'un flexème  $i$  est nul dès lors qu'un flexème n'affiche dans son paradigme qu'une seule réalisation pour la formation de tous les radicaux. En dehors des cas de supplétion au sens strict où les radicaux sont listés explicitement dans le lexique pour une entrée donnée, chaque réalisation radicale distincte est le résultat d'une règle de réalisation provenant d'une zone réalisationnelle différente de celles des autres radicaux à au moins un niveau réalisationnel de type radical ou thématique. Ainsi, si à au moins niveau réalisationnel de type radical  $l^s$  ou thématique  $l^{th}$  un flexème  $i$  est associé par son schème à plus d'une zone, ce flexème affichera un cas d'allomorphie radicale.

Autrement dit, un flexème  $i$  affiche de l'allomorphie radicale régulière si et seulement si il existe parmi les radicaux de son paradigme au moins un  $s_1$  de forme  $(z_1^s, r_1^s)$  et un  $s_2$  de forme  $(z_2^s, r_2^s)$  tel que  $r_1^s$  est différent de  $r_2^s$ .

Nous noterons  $S_{dist}^{reg}(i)$  l'ensemble des radicaux distincts dans la paradigme de  $i$  qui ont été construits par des règles de réalisations régulières. Il correspond à l'ensemble de radicaux distincts  $S_{dist}(i)$  privé de l'ensemble des radicaux lexicalement spécifiés<sup>40</sup>.

40. Nous rappelons qu'un radical  $s$  est défini comme un couple  $(z^s, \text{output})$  où  $z^s$  est une zone de radical

$$S_{dist}(i) = S_{dist}^{reg}(i) \cup S_i^{suppl} \quad (4.17)$$

Sachant que par définition

$$S_{dist}^{reg}(i) \cap S_i^{suppl} = \emptyset \quad (4.18)$$

Nous définissons l'ensemble  $\mathcal{A}_{set}(i)$  des réalisations radicales régulières distinctes tel que pour tout radical  $s = (z^s, r^s)$  appartenant à  $S_{dist}^{reg}(i)$ ,  $r^s$  appartient à  $\mathcal{A}_{set}(i)$ .

$$\mathcal{A}_{set}(i) = \{r^s | (z^s, r^s) \in S_{dist}^{reg}(i)\} \quad (4.19)$$

Une autre manière, plus simple, de dire qu'un flexème  $i$  présente de l'allomorphie radicale est alors de dire que les formes de son paradigme sont construites à partir de plus d'une réalisation radicale. Il remplit ainsi la condition suivante :

$$|\mathcal{A}_{set}(i)| > 1 \quad (4.20)$$

Dans le cas du verbe ASSEoir, les réalisations radicales régulières sont au nombre de cinq, contruites respectivement à partir des couples réalisationnels de niveau radical indiquées dans le schème de l'entrée lexicale d'ASSEoir.

- $(z_{ied}^s, id)$  pour *assied-*,
- $(z_{oi}^s, id)$  pour *assoi-*,
- $(z_{oy}^s, id)$  pour *assoy-*,
- $(z_{ey}^s, id)$  pour *assey-*,
- $(z_i^s, id)$  pour *assi-*.

L'ensemble  $\mathcal{A}_{set}(\text{ASSEoir})$  est ainsi :

---

et  $r^s$  la réalisation phonique du radical.

$$\begin{aligned} \mathcal{A}_{set}(\text{ASSEOIR}) &= \{r^s \mid (z^s, r^s) \in \mathcal{S}_{dist}^{reg}(\text{ASSOIR})\} \\ &= \{\text{assied-}, \text{assoï-}, \text{assey-}, \text{asseï-}, \text{assi-}\} \end{aligned} \quad (4.21)$$

Le cardinal de  $\mathcal{A}_{set}(\text{ASSEOIR})$  vaut ainsi :

$$|\mathcal{A}_{set}(\text{ASSEOIR})| = 5 > 1 \quad (4.22)$$

Il est ainsi supérieur à 1, ce qui fait d'ASSEOIR un verbe présentant de l'allomorphie radicale.

**Indice d'allomorphie radicale d'un flexème donné** L'indice d'allomorphie radicale d'un flexème  $i$  correspond alors au nombre maximal de réalisations de type radical différentes formées de façon régulière pour ce flexème moins 1.

L'indice  $\mathcal{A}(i)$  d'allomorphie radicale de  $i$  est ainsi calculée comme suit :

$$\mathcal{A}(i) = |\mathcal{A}_{set}(i)| - 1 \quad (4.23)$$

L'indice  $\mathcal{A}(\text{ASSEOIR})$  du verbe ASSEOIR vaut par conséquent :

$$\begin{aligned} \mathcal{A}(\text{ASSEOIR}) &= |\mathcal{A}_{set}(\text{ASSEOIR})| - 1 \\ &= 5 - 1 \\ &= 4 \end{aligned} \quad (4.24)$$

**Indice d'allomorphie radicale globale d'un système flexionnel** L'indice d'allomorphie radicale globale d'un système flexionnel est la moyenne des indices d'allomorphie radicale des flexèmes de son lexique flexionnel :

$$\mathcal{A}(\mu_{infl}) = \frac{\sum_{i \in \mathcal{I}} \mathcal{A}(i)}{|\mathcal{I}|} \quad (4.25)$$

**Indice de dispersion de l'allomorphie radicale d'un système morphologique** L'indice de dispersion de l'allomorphie radicale dans un système flexionnel permet de mesurer

l'homogénéité de la répartition des phénomènes d'allomorphie radicale dans le système flexionnel en question. Il correspond au nombre de flexèmes dont l'indice  $\mathcal{A}(i)$  d'allomorphie radicale n'est pas nul divisé par la taille du lexique flexionnel.

L'indice de dispersion  $\mathcal{A}_{disp}(\mu_{infl})$  de l'allomorphie radicale du système flexionnel  $\mu_{infl}$  est calculé comme suit :

$$\mathcal{A}_{disp}(\mu_{infl}) = \frac{|\{i \in \mathcal{S} | \mathcal{A}(i) > 0\}|}{|\mathcal{S}|} \quad (4.26)$$

**Indice d'allomorphie radicale moyenne d'un système flexionnel** Si aucun flexème d'un système flexionnel ne présente d'allomorphie radicale, l'indice d'allomorphie radicale moyenne de ce système flexionnel est définie comme étant nulle. Sinon, l'indice d'allomorphie radicale moyenne de ce système flexionnel est la somme des indices d'allomorphie radicale des flexèmes de son lexique flexionnel divisée par le nombre de flexèmes du lexique flexionnel présentant une allomorphie radicale.

L'indice  $\mathcal{A}_{av}(\mu_{infl})$  d'allomorphie radicale moyenné du système flexionnel  $\mu_{infl}$  est obtenu comme suit :

$$\mathcal{A}_{av}(\mu_{infl}) = \frac{\sum_{i \in \mathcal{S}} \mathcal{A}(i)}{|\{i \in \mathcal{S} | \mathcal{A}(i) > 0\}|} \quad (4.27)$$

#### 4.4.1.4 Indice d'hétéroclité

**Hétéroclise** Un flexème  $i$  est considéré comme hétéroclite si et seulement si, pour deux PTM  $\sigma_1$  et  $\sigma_2$  distincts appartenant à l'ensemble des PTM  $\Sigma_i$  pour lequel se fléchit un flexème  $i$ , il existe un niveau réalisationnel  $l \in \{l_{i,1}, \dots, l_{i,n}\}$  où  $\sigma_1$  est associé à une zone réalisationnelle de d'une classe réalisationnelle  $Z_1$  et  $\sigma_2$  à une zone d'une classe distincte  $Z_2$ .

Nous avons illustré le phénomène de l'hétéroclise par la flexion de certains noms d'animaux du slovaque, notamment à l'exemple du nom OROL « aigle » dont le paradigme figure au tableau 4.3 page 199. Comme le montrent les sous-schémas listés dans le schéma de son entrée lexicale, illustrée par la figure 4.30, OROL se fléchit pour son sous-paradigme du singulier selon la zone réalisationnelle  $Z_{ANIM, SG}^{exp}$  comme un nom masculin animé de la classe flexionnelle  $Z_{anim}^{exp}$  et pour son sous-paradigme du pluriel selon la zone réalisationnelle  $Z_{INAN, PL}^{exp}$  comme un nom masculin inanimé de classe  $Z_{inan}^{exp}$ .

Au niveau exposant, les zones réalisationnelles déterminant la flexion d'OROL appartiennent ainsi à deux classes réalisationnelles différentes. OROL est ainsi un flexème hété-

OROL					
I-PHON	orol				
I-CAT	<i>nom</i>				
MSF	{ <i>standard</i> }				
S-STEM	( <i>vide</i> )				
S-FORM	( <i>vide</i> )				
I-PAT	<table style="border-collapse: collapse; width: 100%;"> <tr> <td style="border-right: 1px solid black; padding: 2px 5px;"><math>(Z_{\text{reg}}^s, \text{id})</math>,</td> <td style="padding: 2px 5px;"><math>(z_{\text{ANIM, SG}}^{\text{exp}}, \text{id})</math></td> </tr> <tr> <td style="border-right: 1px solid black; padding: 2px 5px;"><math>(Z_{\text{reg}}^s, \text{id})</math>,</td> <td style="padding: 2px 5px;"><math>(z_{\text{INAN, PL}}^{\text{exp}}, \text{id})</math></td> </tr> </table>	$(Z_{\text{reg}}^s, \text{id})$ ,	$(z_{\text{ANIM, SG}}^{\text{exp}}, \text{id})$	$(Z_{\text{reg}}^s, \text{id})$ ,	$(z_{\text{INAN, PL}}^{\text{exp}}, \text{id})$
$(Z_{\text{reg}}^s, \text{id})$ ,	$(z_{\text{ANIM, SG}}^{\text{exp}}, \text{id})$				
$(Z_{\text{reg}}^s, \text{id})$ ,	$(z_{\text{INAN, PL}}^{\text{exp}}, \text{id})$				

FIGURE 4.30 : Entrée lexicale pour le nom slovaque OROL

roclite.

**Indice d'hétérocliticité d'un niveau de réalisation pour un flexème donné** L'indice d'hétérocliticité d'un niveau réalisationnel  $l \in \{l_{i,1}, \dots, l_{i,n}\}$  pour un flexème  $i$  donné est défini comme le nombre de classes flexionnelles auxquelles appartiennent les zones réalisationnelles appartenant à la règle de sélection de zones réalisationnelles  $\mathbb{L}_{l|z}^i$ .

Soit  $z$  une zone appartenant à un niveau  $l \in \{l_{i,1}, \dots, l_{i,n}\}$  donné. Nous notons  $\hat{z}$  la classe flexionnelle dans  $l$  à laquelle est par défaut associée  $z$ .

L'indice d'hétérocliticité  $\mathcal{H}_l(i)$  du niveau  $l \in \{l_{i,1}, \dots, l_{i,n}\}$  pour  $i$  est alors défini comme le nombre de classes flexionnelles distinctes auxquelles appartiennent les zones réalisationnelles associées à  $i$  au niveau  $l$  par  $\mathbb{L}_{l|z}^i$  moins 1.

$$\mathcal{H}_l(i) = |\{\hat{z} | z \in \mathbb{L}_{l|z}^i\}| - 1 \quad (4.28)$$

Dans le cas d'OROL, l'indice d'hétérocliticité au niveau exposant vaut ainsi :

$$\begin{aligned} \mathcal{H}_l(\text{OROL}) &= |\{Z_{\text{ANIM}}^{\text{exp}}, Z_{\text{INAN}}^{\text{exp}}\}| - 1 \\ &= 2 - 1 \\ &= 1 \end{aligned} \quad (4.29)$$

Au niveau radical, l'indice d'hétérocliticité d'OROL est nul :



$$\begin{aligned}
 \mathcal{H}_l(\text{OROL}) &= |\{Z_{\text{REG}}^s\}| - 1 \\
 &= 1 - 1 \\
 &= 0
 \end{aligned}
 \tag{4.30}$$

**Indice d'hétérocliticité d'un flexème donné** L'indice d'hétérocliticité  $\mathcal{H}(i)$  d'un flexème  $i$  est défini comme la somme de ses indices d'hétérocliticité par niveau  $\mathcal{H}_l(i)$  pour l'ensemble des niveaux réalisationnels  $\{l_{i,1}, \dots, l_{i,n}\}$  intervenant dans la flexion de  $i$ .

$$\mathcal{H}(i) = \sum_{l \in \{l_{i,1}, \dots, l_{i,n}\}} \mathcal{H}_l(i)
 \tag{4.31}$$

L'indice d'hétérocliticité global d'OROL vaut ainsi :

$$\begin{aligned}
 \mathcal{H}(\text{OROL}) &= \sum_{l \in \{l_{\text{OROL},1}, \dots, l_{\text{OROL},n}\}} \mathcal{H}_l(\text{OROL}) \\
 &= 0 + 1 \\
 &= 1
 \end{aligned}
 \tag{4.32}$$

**Indice d'hétérocliticité d'un système flexionnel donné** L'indice d'hétérocliticité  $\mathcal{H}(\mu_{\text{infl}})$  d'un système flexionnel complet peut alors être défini comme la moyenne des indices d'hétérocliticité de ses flexèmes  $i$  sur son lexique flexionnel  $\mathcal{S}$ .

$$\mathcal{H}(\mu_{\text{infl}}) = \frac{\sum_{i \in \mathcal{S}} \mathcal{H}(i)}{|\mathcal{S}|}
 \tag{4.33}$$

#### 4.4.2 Mesures de la canonicité des réalisations

Le deuxième lot de mesures de canonicité concerne les mesures sur les réalisations des radicaux et des formes dans les paradigmes des flexèmes d'un système flexionnel donné.

##### 4.4.2.1 Indice de supplétion de radicaux

**Supplétion de radicaux** Un flexème  $i$  présente de la supplétion de radicaux s'il existe au moins un PTM pour lequel le radical n'est pas construit par l'application des règles de

formation de radical définies pour sa catégorie flexionnelle  $\mathcal{C}_i$ . Autrement dit, l'ensemble  $S_i^{suppl}$  des radicaux supplétifs listés dans l'entrée lexicale de  $i$  n'est pas l'ensemble vide.

Dans le cas du verbe ALLER dont nous rappelons l'entrée lexicale simplifiée à la figure 4.31,  $S_{ALLER}^{suppl}$  contient les radicaux supplétif  $v-$ ,  $aill-$  et  $ir-$ . La liste  $S_{ALLER}^{suppl}$  n'étant pas vide, ALLER présente de la supplétion de radical.

ALLER	
I-PHON	aller
I-CAT	<i>verbe</i>
MSF	{ <i>standard</i> }
S-STEM	$z_1^s$ : <i>v-</i> $z_3^s$ : <i>aill-</i> $z_4^s$ : <i>i-</i>
S-FORM	( <i>vide</i> )
I-PAT	( $Z_{default}^s$ , id), ( $Z_{aller}^{exp}$ , id)

FIGURE 4.31 : Entrée lexicale pour le verbe ALLER

**Indice de supplétion de radicaux pour un flexème donné** L'indice de supplétion de radicaux d'un flexème  $i$  donné est défini comme le nombre de radicaux pour la réalisation desquels les niveaux réalisationnels ont été court-circuités par un radical listé indépendamment dans l'entrée de  $i \in \mathcal{I}$ , c'est à dire le nombre d'éléments dans la liste de radicaux supplétifs indiquée dans l'entrée lexicale de  $i$ .

L'indice de supplétion de radicaux  $\mathcal{R}^s(i)$  d'un flexème  $i$  est ainsi défini comme le cardinal de l'ensemble  $S_i^{suppl}$ .

$$\mathcal{R}^s(i) = |S_i^{suppl}| \tag{4.34}$$

Pour le cas du verbe ALLER, l'indice de supplétion de radicaux vaut ainsi :

$$\begin{aligned} \mathcal{R}^s(ALLER) &= |S_{ALLER}^{suppl}| \\ &= 3 \end{aligned} \tag{4.35}$$

**Indice de supplétion de radicaux pour un système flexionnel donné** L'indice de supplétion de radicaux  $\mathcal{R}^s(\mu_{\text{infl}})$  d'un système flexionnel  $\mu_{\text{infl}}$  est défini comme la moyenne des indices de supplétion de radicaux de ses flexèmes sur son lexique flexionnel.

$$\mathcal{R}^s(\mu_{\text{infl}}) = \frac{\sum_{i \in \mathcal{I}} \mathcal{R}^s(i)}{|\mathcal{I}|} \quad (4.36)$$

#### 4.4.2.2 Indice de supplétion de forme

**Supplétion de forme** Un flexème  $i$  présente de la supplétion de forme s'il existe au moins un PTM pour lequel une forme n'est pas construite par l'application des règles réalisatinnelles définies pour sa catégorie flexionnelle  $\mathcal{C}_i$ . Autrement dit, l'ensemble  $F_i^{\text{suppl}}$  des formes supplétives listées dans l'entrée lexicale de  $i$  n'est pas l'ensemble vide.

Nous avons donné un exemple de supplétion de forme dans le paradigme du verbe ÊTRE dont nous rappelons l'entrée lexicale simplifiée à la figure 4.32. L'ensemble de formes supplétives  $F_{\text{ÊTRE}}^{\text{suppl}}$  comporte trois formes supplétives (*sommes*, 1SG.PRS.IND), (*suis*, 1PL.PRS.IND) et (*sont*, 3PL.PRS.IND). Il n'est donc pas vide, ce qui fait du verbe ÊTRE un verbe présentant de la supplétion de forme.

ÊTRE	
I-PHON	être
I-CAT	<i>verbe</i>
MSF	{ <i>standard</i> }
S-STEM	{ ( <i>vide</i> ) }
S-FORM	1SG.PRS.IND <i>sommes</i> 1PL.PRS.IND <i>suis</i> 3PL.PRS.IND <i>sont</i>
I-PAT	( $Z_{\text{default}}^s, \text{id}$ ), ( $Z_{\text{tre}}^{\text{exp}}, \text{id}$ )

FIGURE 4.32 : Entrée lexicale pour le verbe ÊTRE

**Indice de supplétion de forme pour un flexème donné** L'indice de supplétion de forme d'un flexème  $i$  donné est défini comme le nombre de formes pour la réalisation desquelles les niveaux réalisatinnelles ont été court-circuités par une forme listée

indépendamment dans l'entrée de  $i \in \mathcal{I}$ , c'est à dire le nombre d'éléments dans la liste de formes supplétives indiquée dans l'entrée lexicale de  $i$ .

L'indice de supplétion de formes  $\mathcal{R}^f(i)$  d'un flexème  $i$  est ainsi défini comme le cardinal de l'ensemble  $F_i^{suppl}$ .

$$\mathcal{R}^f(i) = |F_i^{suppl}| \quad (4.37)$$

Pour le cas du verbe ÊTRE, l'indice de supplétion de formes vaut ainsi :

$$\begin{aligned} \mathcal{R}^f(\hat{\text{ÊTRE}}) &= |F_{\hat{\text{ÊTRE}}}^{suppl}| \\ &= 3 \end{aligned} \quad (4.38)$$

**Indice de supplétion de forme pour un système flexionnel donné** L'indice de supplétion de formes  $\mathcal{R}^f(\mu_{\text{infl}})$  d'un système flexionnel  $\mu_{\text{infl}}$  est défini comme la moyenne des indices de supplétion de formes de ses flexèmes sur son lexique flexionnel.

$$\mathcal{R}^f(\mu_{\text{infl}}) = \frac{\sum_{i \in \mathcal{I}} \mathcal{R}^f(i)}{|\mathcal{I}|} \quad (4.39)$$

#### 4.4.2.3 Indice de syncrétisme

Contrairement aux radicaux dont la réalisation ne varie canoniquement pas à travers le paradigme, les formes dans les cases dudit paradigme sont canoniquement toutes différentes. L'indice de syncrétisme est en ce sens une mesure inverse de celle de l'allomorphie radicale, mais appliquée cette fois à la forme complète et non à la seule réalisation des radicaux.

**Présence de syncrétismes dans un paradigme** Le paradigme d'un flexème  $i$  comporte des formes syncrétiques si et seulement s'il existe au moins deux PTM distincts  $\sigma_1$  et  $\sigma_2$ , tel que la réalisation de la forme  $f_1$  exprimant  $\sigma_1$  est identique à la réalisation de la forme  $f_2$  exprimant  $\sigma_2$ .

Par exemple dans les paradigmes des noms féminins en *-a* du latin, le génitif singulier et le datif singulier sont réalisés par la même forme. De même, pour les noms neutres en latin le nominatif et l'accusatif (qu'ils soient singulier ou pluriel) sont toujours syncrétiques. Nous

	ROSA « rose »		DONUM « don »	
	SG	PL	SG	PL
NOM	<i>rosa</i>	<i>rosæ</i>	<i>donum</i>	<i>dona</i>
ACC	<i>rosam</i>	<i>rosas</i>	<i>donum</i>	<i>dona</i>
GEN	<i>rosæ</i>	<i>rosārum</i>	<i>doni</i>	<i>donōrum</i>
DAT	<i>rosæ</i>	<i>rosīs</i>	<i>dono</i>	<i>donīs</i>
ABL	<i>rosā</i>	<i>rosīs</i>	<i>dono</i>	<i>donīs</i>

TABLE 4.19 : Syncrétismes dans les paradigmes de noms latins

donnons une illustration des paradigmes de noms latins féminins en *-a* et neutres en *-um* au tableau 4.19 pour les noms ROSA « rose » et DONUM « don » respectivement. Pour chaque paradigme, les cases syncrétiques sont indiquées par la même teinte dans la case du tableau. Le fait qu'il existe dans ces paradigmes des cases identiques fait des flexèmes auxquels ils appartiennent des flexèmes affichant des cas de syncrétisme.

**Ensemble de cases syncrétiques pour une forme donnée** Un ensemble de cases syncrétiques dans un paradigme est un ensemble de cases comportant les mêmes réalisations.

Soit  $P_i$  le paradigme d'un flexème  $i$  constitué de formes de la forme  $f = (\sigma, r^f)$ . Un ensemble de cases syncrétiques  $\mathcal{S}_F(i, r^f)$  est calculé comme l'ensemble des cases pour lesquels la réalisation formelle  $r^f$  est la même d'une case à l'autre mais le PTM  $\sigma$  varie :

$$\mathcal{S}_F(i, r^f) = \{\sigma \in \Sigma_i \mid (\sigma, r^f) \in P_i\} \quad (4.40)$$

Nous indiquons ainsi les ensembles de cases syncrétiques pour les flexèmes ROSA et DONUM au tableau 4.20.

ROSA « rose »		DOMUM « don »	
$\mathcal{S}_F(\text{ROSA}, \textit{rosa})$	{NOM.SG}	$\mathcal{S}_F(\text{DONUM}, \textit{donum})$	{NOM.SG, ACC.SG}
$\mathcal{S}_F(\text{ROSA}, \textit{rosæ})$	{GEN.SG, DAT.SG, NOM.PL}	$\mathcal{S}_F(\text{DONUM}, \textit{doni})$	{GEN.SG}
$\mathcal{S}_F(\text{ROSA}, \textit{rosam})$	{ACC.SG}	$\mathcal{S}_F(\text{DONUM}, \textit{dono})$	{DAT.SG, ABL.SG}
$\mathcal{S}_F(\text{ROSA}, \textit{rosā})$	{ABL.SG}	$\mathcal{S}_F(\text{DONUM}, \textit{dona})$	{NOM.PL, ACC.PL}
$\mathcal{S}_F(\text{ROSA}, \textit{rosārum})$	{GEN.PL}	$\mathcal{S}_F(\text{DONUM}, \textit{donōrum})$	{GEN.PL}
$\mathcal{S}_F(\text{ROSA}, \textit{rosīs})$	{DAT.PL, ABL.PL}	$\mathcal{S}_F(\text{DONUM}, \textit{donīs})$	{DAT.PL, ABL.PL}

TABLE 4.20 : Ensembles de cases syncrétiques dans les paradigmes de noms latins

**Indice de synchrétisme pour un flexème donné** L'indice de synchrétisme  $\mathcal{S}(i)$  d'un flexème donné se déduit de la somme des tailles des ensembles  $\mathcal{S}_F(i, r^f)$  de cases synchrétiques définis ci-dessus, chaque ensemble de cases synchrétiques comptant pour sa taille moins 1 dans l'indice.

$$\mathcal{S}(i) = \sum_{r^f \in F_{dist}(i)} \left( |\mathcal{S}_F(i, r^f)| - 1 \right) \quad (4.41)$$

Ainsi les indices de synchrétisme  $\mathcal{S}(\text{ROSA})$  et  $\mathcal{S}(\text{DONUM})$  respectivement pour les flexèmes ROSA et DONUM valent-ils :

$$\begin{aligned} \mathcal{S}(\text{ROSA}) &= \sum_{r^f \in F_{dist}(\text{ROSA})} \left( |\mathcal{S}_F(\text{ROSA}, r^f)| - 1 \right) \\ &= (1 - 1) + (3 - 1) + (1 - 1) + (1 - 1) + (1 - 1) + (2 - 1) \\ &= 3 \end{aligned} \quad (4.42)$$

$$\begin{aligned} \mathcal{S}(\text{DONUM}) &= \sum_{r^f \in F_{dist}(\text{DONUM})} \left( |\mathcal{S}_F(\text{DONUM}, r^f)| - 1 \right) \\ &= (2 - 1) + (1 - 1) + (2 - 1) + (2 - 1) + (1 - 1) + (2 - 1) \\ &= 4 \end{aligned} \quad (4.43)$$

**Indice de synchrétisme pour un système flexionnel donné** L'indice de synchrétisme pour un système flexionnel  $\mu_{infl}$  complet est défini comme la moyenne de l'indice de synchrétisme  $\mathcal{S}(i)$  de tous les lexèmes sur l'ensemble du lexique flexionnel.

$$\mathcal{S}(\mu_{infl}) = \frac{\sum_{i \in \mathcal{I}} \mathcal{S}(i)}{|\mathcal{I}|} \quad (4.44)$$

### 4.4.3 Mesures de réalisation canoniques des traits morphosyntaxiques

Le dernier ensemble de mesures de canonicité évalue le degré de canonicité de la réalisation des PTM pour lesquels un flexème  $i$  d'un système flexionnel  $\mu_{infl}$  est défini. Ces mesures quantifient ainsi les cas de déficience, de surdifférenciation et de décalage

morphosyntaxique.

#### 4.4.3.1 Indice de déficience

**Déficience d'un flexème** Un flexème  $i$  sera considéré comme déficient s'il ne réalise pas l'intégralité des PTM canoniquement réalisés par sa catégorie flexionnelle  $\mathcal{C}_i$ . Il sera ainsi déficient si et seulement s'il y a dans l'ensemble  $\Sigma_{\mathcal{C}_i}$  des PTM réalisés par les flexèmes appartenant canoniquement à la catégorie flexionnelle  $\mathcal{C}_i$  qui ne sont pas dans l'ensemble  $\Sigma_i$  des PTM réalisés par  $i$ .

D'après l'entrée lexicale de SEOIR que nous rappelons à la figure 4.33, le flexème SEOIR ne réalise par exemple pas de forme pour la première personne du singulier de l'indicatif présent. Le PTM {1SG.PRS.IND} n'appartient pas à l'ensemble  $\Sigma_{\text{seoir}}$  des PTM réalisés par SEOIR alors même qu'il appartient à l'ensemble  $\Sigma_{\text{VERBE}}$  des PTM réalisés par les membres canoniques de la catégorie flexionnelle VERBE en français. Le verbe SEOIR est donc déficient.

SEOIR	
I-PHON	seoir
I-CAT	<i>verbe</i>
MSF	$\left\{ \begin{array}{l} 3\text{SG.PRS.IND}, 3\text{PL.PRS.IND}, 3\text{SG.PST.IND}, \\ 3\text{PL.PST.IND}, 3\text{SG.FUT.IND}, 3\text{PL.FUT.IND}, \\ 3\text{SG.PRS.COND}, 3\text{PL.PRS.COND}, 3\text{SG.PRS.SBJV}, \\ 3\text{PL.PRS.SBJV}, \text{INF}, \text{PTCP} \end{array} \right\}$
S-STEM	( <i>vide</i> )
S-FORM	( <i>vide</i> )
I-PAT	$(z_{ie}^s, \text{id}), (Z_{\text{SEOIR}}^{\text{exp}}, \text{id})$ $(z_{ey}^s, \text{id}), (Z_{\text{SEOIR}}^{\text{exp}}, \text{id})$ $(z_e^s, \text{id}), (Z_{\text{SEOIR}}^{\text{exp}}, \text{id})$

FIGURE 4.33 : Entrée lexicale pour le verbe SEOIR

**Indice de déficience d'un paradigme donné** L'indice de déficience d'un paradigme donné est défini comme le nombre de PTM compris dans  $\Sigma_{\mathcal{C}_i}$  qui ne sont pas compris dans  $\Sigma_i$ .

$$\mathcal{D}^-(i) = |\Sigma_{\mathcal{C}_i} \setminus \Sigma_i| \quad (4.45)$$

Le verbe *SEOIR* ne comporte ainsi dans son inventaire  $\Sigma_{\text{SEOIR}}$  de PTM que 13 des 48 PTM canoniquement réalisés synthétiquement par les flexèmes verbaux (Arrivé, 1997). Son indice de déficience vaut ainsi :

$$\begin{aligned} \mathcal{D}^-(\text{SEOIR}) &= |\Sigma_{\text{VERBE}} \setminus \Sigma_{\text{SEOIR}}| \\ &= 48 - 13 \\ &= 35 \end{aligned} \quad (4.46)$$

**Indice de déficience d'un système flexionnel donné** L'indice de déficience  $\mathcal{D}^-(\mu_{\text{infl}})$  d'un système flexionnel  $\mu_{\text{infl}}$  est la moyenne des indices de déficience de ses flexèmes  $\mathcal{D}^-(i)$  sur le lexique flexionnel.

$$\mathcal{D}^-(\mu_{\text{infl}}) = \frac{\sum_{i \in \mathcal{I}} \mathcal{D}^-(i)}{|\mathcal{I}|} \quad (4.47)$$

#### 4.4.3.2 Indice de surdifférenciation

**Surdifférenciation dans le paradigme d'un flexème** Un flexème  $i$  sera considéré comme affichant de la surdifférenciation s'il réalise des PTM qui ne sont pas réalisés par les membres canoniques de sa catégorie flexionnelle  $\mathcal{C}_i$ . Il sera ainsi considéré comme affichant de la surdifférenciation si et seulement s'il y a dans l'ensemble  $\Sigma_i$  des PTM réalisés par  $i$  qui ne sont pas dans l'ensemble  $\Sigma_{\mathcal{C}_i}$  des PTM canoniquement réalisés par les flexèmes appartenant à la catégorie flexionnelle  $\mathcal{C}_i$ .

$$\Sigma_i \setminus \Sigma_{\mathcal{C}_i} \neq \emptyset \quad (4.48)$$

Un exemple que nous avons donné plus haut, au tableau 4.13 page 224, est celui du vocatif en slovaque. Seuls quelques rares noms comme *PRIATEL* distinguent dans leurs formes le PTM du vocatif singulier. Nous rappelons l'entrée lexicale d'un tel nom, le nom *PRIATEL*, à la figure 4.34. Cette entrée présente un PTM *VOC.SG* surnuméraire par rapport



aux douze PTM habituellement réalisés par les noms du slovaque. Il présente donc de la surdifférenciation.

PRIATEL	
I-PHON	priatel̥
I-CAT	<i>nom</i>
MSF	$\left\{ \begin{array}{l} \text{NOM.SG, GEN.SG, DAT.SG, ACC.SG,} \\ \text{VOC.SG, LOC.SG, INS.SG,} \\ \text{NOM.PL, GEN.PL, DAT.PL, ACC.PL,} \\ \text{LOC.PL, INS.PL} \end{array} \right\}$
S-STEM	( <i>vide</i> )
S-FORM	( <i>vide</i> )
I-PAT	$\left( \begin{array}{l} (z_C^s, \text{id}), (z_{M.ANIM-C}^{exp}, \text{id}) \\ (z_C^s, \text{id}), (z_{VOC}^{exp}, \text{id}) \end{array} \right)$

FIGURE 4.34 : Entrée lexicale pour le nom slovaque PRIATEL

**Indice de surdifférenciation d'un paradigme donné** L'indice de surdifférenciation d'un paradigme donné est défini comme le nombre de PTM compris dans  $\Sigma_i$  qui ne sont pas compris dans  $\Sigma_{e_i}$ .

L'indice de déficience d'un paradigme  $\mathcal{D}^+(i)$  est ainsi calculé comme suit :

$$\mathcal{D}^+(i) = |\Sigma_i \setminus \Sigma_{e_i}| \quad (4.49)$$

Dans le cas de PRIATEL, l'indice de surdifférenciation de son paradigme vaut ainsi :

$$\begin{aligned} \mathcal{D}^+(\text{PRIATEL}) &= |\Sigma_{\text{PRIATEL}} \setminus \Sigma_{\text{NOM}}| \\ &= 13 - 12 \\ &= 1 \end{aligned} \quad (4.50)$$

**Indice de surdifférenciation d'un système flexionnel donné** L'indice de surdifférenciation  $\mathcal{D}^+(\mu_{\text{infl}})$  d'un système flexionnel  $\mu_{\text{infl}}$  est la moyenne des indices de surdifférenciation de ses flexèmes  $\mathcal{D}^+(i)$  sur l'ensemble du lexique flexionnel.

$$\mathcal{D}^+(\mu_{\text{infl}}) = \frac{\sum_{i \in \mathcal{I}} \mathcal{D}^+(i)}{|\mathcal{I}|} \quad (4.51)$$

#### 4.4.3.3 Indice de décalage morphosyntaxique

**Décalage morphosyntaxique** Les phénomènes de décalage morphosyntaxique sont les cas où, pour la réalisation d'une forme d'un flexème  $i$  donné, un trait  $\sigma' \in t(\Sigma_i)$  réalisé par une règle de réalisation  $R$  donnée n'est pas le même que celui pour lequel la zone de sélection de règles de réalisation  $z$  à laquelle appartient  $R$  avait été sélectionnée. Autrement dit, ce sont les zones réalisationnelles de type  $(z, t)$  pour lesquelles  $t \neq \text{id}$ .

Un flexème  $i$  sera ainsi considéré comme décalé si et seulement s'il existe un PTM  $\sigma \in \Sigma_i$  tel que dans au moins un des sous-schémas du schéma  $\mathbb{P}_i$  de  $i$  figure au moins un couple réalisationnel  $(z, t)$  tel que  $t$  n'est pas la fonction identité  $\text{id}$ .

Un exemple de cela est le nom serbo-croate DETE « *enfant* » que nous avons discuté plus haut et dont nous rappelons l'entrée lexicale à la figure 4.35. L'entrée indique dans le schéma flexionnel  $\mathbb{P}_{\text{DETE}}$  de DETE le couple réalisationnel  $(z_{F-a,sg}^{\text{exp}}, t_{\text{NB}})$  où la fonction de transfert  $t_{\text{NB}}$  n'est pas la fonction identité. DETE présente ainsi un décalage morphosyntaxique.

DETE	
I-PHON	dete
I-CAT	<i>nom</i>
MSF	{ <i>standard</i> }
S-STEM	( <i>vide</i> )
S-FORM	( <i>vide</i> )
I-PAT	( $Z_{\text{reg}}^s, \text{id}$ ), ( $z_{\text{N-ET, SG}}^{\text{exp}}, \text{id}$ )
	( $Z_{\text{reg}}^s, \text{id}$ ), ( $z_{\text{F-A, SG}}^{\text{exp}}, t_{\text{NB}}$ )

FIGURE 4.35 : Entrée lexicale pour le nom serbo-croate DETE « *enfant* »

**Indice de décalage morphosyntaxique d'un flexème pour un niveau donné** Pour définir l'indice de décalage morphosyntaxique d'un flexème pour un niveau donné, nous

avons besoin de définir  $\mathcal{P}_{\Sigma_i, l}$ , la partition<sup>41</sup> optimale de niveau  $l$  de  $\Sigma_i$  selon  $\mathbb{L}_l^i$ .  $\mathcal{P}_{\Sigma_i, l}$  est définie de telle façon que pour tout ensemble  $\Sigma \in \mathcal{P}_{\Sigma_i, l}$ , tous les PTM  $\sigma \in \Sigma$  sont associés à la même fonction de transfert  $t_{\Sigma, l}$  par  $\mathbb{L}_l^i$ ; de plus, pour tout  $\Sigma' \in \mathcal{P}_{\Sigma_i, l}$  tel que  $\Sigma' \neq \Sigma$ ,  $\sigma' \in \Sigma'$  est associée par  $\mathbb{L}_l^i$  à une fonction de transfert  $t \neq t_{\Sigma, l}$ . En d'autres termes, deux PTM ont la même fonction de transfert si et seulement s'ils appartiennent au même élément de la partition  $\mathcal{P}_{\Sigma_i, l}$ .

L'indice de décalage morphosyntaxique  $\mathcal{M}_l(i)$  d'un flexème pour un niveau  $l \in \{l_{i,1}, \dots, l_{i,n}\}$  donné est alors défini comme le nombre d'ensembles de PTM  $\Sigma$  appartenant à la partition de  $\mathcal{P}_{\Sigma_i, l}$  de  $\Sigma_i$  tel que la règle de transfert  $t_{\Sigma, l}$  associée au niveau  $l$  à tous les éléments de  $\Sigma$  n'est pas la fonction identité (notée  $\text{id}$ ).

$$\mathcal{M}_l(i) = |\{\Sigma \in \mathcal{P}_{\Sigma_i, l} \mid \forall \sigma \in \Sigma, (z_{\sigma, l}, t_{\Sigma, l}) \in \mathbb{L}_l^i(\Sigma_i) \Rightarrow t_{\Sigma, l} \neq \text{id}\}| \quad (4.52)$$

Pour un niveau  $l \in \{l_{i,1}, \dots, l_{i,n}\}$ , un flexème est ainsi considéré comme étant

- non décalé si et seulement si  $\mathcal{M}_l(i) = 0$
- et décalé si et seulement si  $\mathcal{M}_l(i) > 0$ ,

Nous dirons plus précisément qu'au niveau  $l$  appartenant à l'ensemble des niveaux  $\{l_{i,1}, \dots, l_{i,n}\}$  définis pour la catégorie morphologique de  $i$ ,  $i$  est

- semi-décalé si et seulement si  $|\mathcal{P}_{\Sigma_i, l}| > \mathcal{M}_l(i) > 0$
- complètement décalé si et seulement si  $\mathcal{M}_l(i) = |\mathcal{P}_{\Sigma_i, l}|$ .

Dans le cas de  $\text{DETE}$ , nous pourrons ainsi calculer les deux indices  $\mathcal{M}_{l^s}(\text{DETE})$  et  $\mathcal{M}_{l^{\text{exp}}}(\text{DETE})$ , respectivement pour le niveau radical  $l^s$ , et le niveau exposant  $l^{\text{exp}}$  de la façon suivante :

$$\begin{aligned} \mathcal{M}_{l^s}(\text{DETE}) &= |\{\Sigma \in \mathcal{P}_{\Sigma_{\text{DETE}}, l^s} \mid \forall \sigma \in \Sigma, (z_{\sigma, l^s}, t_{\Sigma, l^s}) \in \mathbb{L}_{\text{DETE}}^{l^s}(\Sigma_{\text{DETE}}) \Rightarrow t_{\Sigma, l^s} \neq \text{id}\}| \\ &= 0 \end{aligned} \quad (4.53)$$

$$\begin{aligned} \mathcal{M}_{l^{\text{exp}}}(\text{DETE}) &= |\{\Sigma \in \mathcal{P}_{\Sigma_{\text{DETE}}, l^{\text{exp}}} \mid \forall \sigma \in \Sigma, (z_{\sigma, l^{\text{exp}}}, t_{\Sigma, l^{\text{exp}}}) \in \mathbb{L}_{\text{DETE}}^{l^{\text{exp}}}(\Sigma_{\text{DETE}}) \Rightarrow t_{\Sigma, l^{\text{exp}}} \neq \text{id}\}| \\ &= 1 \end{aligned} \quad (4.54)$$

41. Nous rappelons que la définition d'une PARTITION D'UN ENSEMBLE  $E$  est l'ensemble des sous-ensembles de  $E$  disjoints deux-à-deux et dont l'union est égale à  $E$ .

Au niveau radical, DETE est ainsi non décalé. Au niveau exposant, les deux couples réalisationnels du schème flexionnels présentent des fonctions de transferts différentes. Le nombre de partitions de  $\Sigma_{\text{DETE}}$  au niveau exposant est ainsi de 2, de sorte que DETE est ainsi semi-décalé au niveau exposant.

**Indice de décalage morphosyntaxique global pour un flexème donné** Pour définir l'indice de décalage morphosyntaxique global d'un flexème  $i$ , nous définissons de façon analogue  $\mathcal{P}_{\Sigma_i}$ , la partition optimale de  $\Sigma_i$  selon l'ensemble  $\{\mathbb{L}_{l_1}^i, \dots, \mathbb{L}_{l_n}^i\}$  des règles de sélection de couples réalisationnels définies pour  $i$ .  $\mathcal{P}_{\Sigma_i}$  est définie de telle façon que pour tout ensemble  $\Sigma \in \mathcal{P}_{\Sigma_i}$ , tous les PTM  $\sigma \in \Sigma$  sont associés à chaque niveau  $l \in \{l_{e_i,1}, \dots, l_{e_i,n}\}$  à la même fonction de transfert par  $\mathbb{L}_l^i$ ; de plus, pour tout  $\Sigma' \in \mathcal{P}_{\Sigma_i}$  tel que  $\Sigma' \neq \Sigma$ , il existe au moins un niveau  $l_i \in \{l_{e_i,1}, \dots, l_{e_i,n}\}$  auquel  $\mathbb{L}_{l_i}^i$  n'associe pas la même fonction de transfert à  $\sigma$  et à  $\sigma'$ .

En d'autres termes, deux PTM ont le même ensemble de fonctions de transfert, tous niveaux confondus, si et seulement s'ils appartiennent au même élément de la partition  $\mathcal{P}_{\Sigma_i}$ .

L'indice de décalage morphosyntaxique global  $\mathcal{M}(i)$  d'un flexème  $i$  est alors défini comme le nombre d'ensembles de PTM  $\Sigma$  appartenant à la partition  $\mathcal{P}_{\Sigma_i}$ , tel qu'il existe au moins un niveau  $l_i \in \{l_{e_i,1}, \dots, l_{e_i,n}\}$  auquel la règle de transfert  $t_{\Sigma, l_i}$ , nécessairement partagée par tout  $\sigma \in \Sigma$ , n'est pas la fonction identité  $\text{id}$ .

$$\mathcal{M}(i) = |\{\Sigma \in \mathcal{P}_{\Sigma_i} | (\exists l_{\Sigma} \in \{l_{e_i,1}, \dots, l_{e_i,n}\}) | \forall \sigma \in \Sigma, (z_{\sigma, l_{\Sigma}}, t_{\sigma, l_{\Sigma}}) \in \mathbb{L}_{l_{\Sigma}}^i(\Sigma_i) \Rightarrow t_{\sigma, l_{\Sigma}} \neq \text{id})\}|, \quad (4.55)$$

où l'on rappelle que si  $\sigma$  et  $\sigma'$  sont deux éléments de  $\Sigma \in \mathcal{P}_{\Sigma_i}$ , alors ils partagent à chaque niveau une même fonction de transfert. Autrement dit, dans la formule ci-dessus, tous les  $t_{\sigma, l_{\Sigma}}$  correspondant à des  $\sigma$  d'un même  $\Sigma \in \mathcal{P}_{\Sigma_i}$  sont identiques.

Ainsi, un paradigme est considéré comme étant

- non décalé si et seulement si  $\mathcal{M}(i) = 0$
- décalé si et seulement si  $\mathcal{M}(i) > 0$ , dont
  - semi-décalé si et seulement si  $|\mathcal{P}_{\Sigma_i}| > \mathcal{M}(i) > 0$
  - complètement décalé si et seulement si  $\mathcal{M}(i) = |\mathcal{P}_{\Sigma_i}|$

Dans le cas de DETE indiqué ci-dessus, le seul niveau où il existe deux fonctions de transfert distinctes est le niveau exposant. La partition  $\mathcal{P}_{\Sigma_{\text{DETE}}}$  selon la définition ci-dessus

pour tous les niveaux est donc identique à celle calculée pour le seul niveau exposant plus haut et contient ainsi deux éléments. Il s'ensuit que l'indice de décalage morphosyntaxique  $\mathcal{M}(\text{DETE})$  vaut :

$$\begin{aligned} \mathcal{M}(\text{DETE}) &= |\{\Sigma \in \mathcal{P}_{\Sigma_{\text{DETE}}} | (\exists l_{\Sigma} \in \{l_{e_{\text{NOM}},1}, \dots, l_{e_{\text{NOM}},n}\} | \\ &\quad \forall \sigma \in \Sigma, (z_{\sigma,l_{\Sigma}}, t_{\sigma,l_{\Sigma}}) \in \mathbb{L}_{l_{\Sigma}}^{\text{DETE}}(\Sigma_{\text{DETE}}) \Rightarrow t_{\sigma,l_{\Sigma}} \neq \text{id})\}| \quad (4.56) \\ &= 1 \end{aligned}$$

**Indice de décalage morphosyntaxique pour un système flexionnel donné** À partir de ces définitions, l'indice de décalage morphosyntaxique  $\mathcal{M}(\mu_{\text{infl}})$  d'un système flexionnel  $\mu_{\text{infl}}$  est défini comme la moyenne des indices de décalage morphosyntaxique de chacun de ses flexèmes sur l'ensemble du lexique flexionnel.

$$\mathcal{M}(\mu_{\text{infl}}) = \frac{\sum_{i \in \mathcal{I}} \mathcal{M}(i)}{|\mathcal{I}|} \quad (4.57)$$

## 4.5 Résumé et discussion

Dans ce chapitre, avons introduit notre modèle de la morphologie flexionnelle  $\mathbb{P}\mathbb{R}\mathbb{S}\mathbb{L}$  ainsi que les mesures de canonicité définies à l'intérieur de ce modèle. Ces mesures nous permettent de quantifier le degré de non-canonicité d'un système flexionnel dans une description donnée de ce système pour un phénomène donné et de dresser un portrait global du type de déviation par rapport au canon dont ce système fait preuve.

Mais à ce stade, nous n'avons pas encore résolu un problème supplémentaire qui se pose dès lors qu'un système peut être décrit (de façon également exhaustive) de deux manières différentes qui cependant comportent des degrés de canonicité divers. Prenons par exemple le cas de l'hétéroclise en slovaque. Nous avons décrit la flexion des noms d'animaux slovaques comme relevant de l'emploi hétéroclite de deux zones flexionnelles appartenant respectivement pour le singulier à la classe flexionnelle des masculins animés et pour le pluriel à celle des masculins inanimés. Il aurait en principe également été possible de poser l'existence d'une troisième classe flexionnelle regroupant justement ces deux zones. Le degré d'hétérocliticité des paradigmes concernés par ces zones, et par là celui l'intégralité du système flexionnel du slovaque, dépend ainsi directement des classes réalisationnelles définies dans la description. Il nous manque ainsi un critère

d'évaluation externe des descriptions possibles permettant de choisir de façon objective entre des descriptions concurrentes<sup>42</sup>. Ce n'est qu'une fois que nous aurons ainsi évalué les différentes descriptions possibles que nous pourrions mesurer de façon cohérente le degré de canonicité pour la description effectivement retenue.

Au chapitre 6, nous proposerons en la motivant la COMPACTÉ DE DESCRIPTION comme mesure d'évaluation externe adéquate pour la comparaison de description morphologiques concurrentes. L'applicabilité de notre mesure de compacité repose néanmoins sur une implémentation des descriptions PARSLI dont nous décrivons le mode au chapitre 5.

---

42. C'est-à-dire rendant compte correctement de la quasi-exhaustivité des données flexionnelles d'un système.



## Implémentation de descriptions

Au chapitre 4, nous avons présenté le modèle formel de la morphologie flexionnelle  $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$ . Dans ce chapitre, nous nous intéressons plus en détail à la question de l'implémentation de descriptions morphologiques flexionnelles. En particulier, nous discutons en un premier temps l'importance de l'implémentation pour l'analyse en morphologie flexionnelle. Nous y étudions notamment le caractère plus ou moins explicite dont font preuve les différentes approches formelles de la morphologie, en nous concentrant ici exclusivement sur la représentation de la morphologie flexionnelle. Nous décrivons ensuite le système mis en place pour implémenter spécifiquement des analyses réalisées au sein du modèle  $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$ . Ce système d'implémentation, *Alexina $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$* , s'appuie sur une version modifiée du formalisme lexical *Alexina* (Sagot, 2010). À la section 5.3, nous décrivons enfin une analyse  $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$  de la flexion verbale du maltais et son implémentation en *Alexina $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$*  Sagot & Walther (2013).

### 5.1 Pourquoi implémenter ?

Contrairement à la syntaxe et à la morphologie dérivationnelle, la morphologie flexionnelle présente l'avantage de traiter, pour chaque langue, d'un domaine d'étude fini. Pour un ensemble d'unités lexicales défini, il est possible de lister et de décrire l'intégralité des formes. On est ainsi en droit d'attendre d'une approche formelle de la morphologie qu'elle permette de rendre compte d'un inventaire exhaustif des formes fléchies d'un ensemble fini (mais grand) de lexèmes d'une langue. Pour cela, de nombreuses solutions d'implémentation des analyses formelles existent. Nous défendons ainsi le point de vue que dans la mesure où cela ne pose de problèmes ni techniques ni théoriques, implémenter



une analyse formelle en morphologie flexionnelle apparaît comme essentiel pour tester couverture et la cohérence interne de l'analyse proposée.

### 5.1.1 Approches théoriques et implémentations

Les modèles formels théoriques proposés sont rarement explicites à cet égard. Les formalisations sont souvent assez approximatives, et ne concernent que la représentation d'un phénomène ponctuel souvent extrait du contexte général dans lequel il se situe au sein de la langue étudiée<sup>1</sup>. Peu de modèles existent notamment qui s'accompagnent de véritables implémentations permettant de vérifier les assertions théoriques posées. Parmi les exceptions à cet égard figurent les approches PFM (Stump, 2001) et *Network Morphology* (Brown & Hippisley, 2012)<sup>2</sup>. Chacune de ces approches est accompagnée de moyens d'implémentation adaptés : le formalisme DATR pour *Network Morphology* et les implémentations *Cat's Claw* développées par Raphael Finkel pour PFM<sup>3</sup>.

Cela dit, mêmes les rares exceptions à cet égard ne traitent le plus souvent dans leurs implémentations qu'une quantité limitée de données. Comme le montrent les analyses accessibles en ligne, les implémentations de PFM ne concernent en général qu'entre une entrée et une cinquantaine d'entrées par langue dans les implémentations accessibles sur le site de *Cat's Claw*. Une exception importante est l'analyse du russe de Brown & Hippisley (2012) accessible sur le site du livre de référence de *Network Morphology*<sup>4</sup> : elle comprend un lexique de 1 500 entrées nominales.

Pourtant une implémentation à grande échelle augmente la pertinence d'une analyse dans la mesure où elle permet d'une part de s'assurer que certains phénomènes marginaux n'ont pas été omis dans l'analyse — sans pour autant surévaluer l'importance desdits phénomènes en leur donnant un statut trop central dans l'analyse — et d'autre part de donner à l'inverse toute leur importance aux phénomènes centraux de la langue.

### 5.1.2 Les approches computationnelles

Les approches computationnelles de la morphologie ont pour leur part l'avantage d'être totalement explicites quant à la réalisation des formes.

En dehors des approches qui s'appuient sur des langages de descriptions de réseaux

---

1. Voir par exemple notre critique du traitement des déponents latins par Kiparsky (2005) au chapitre 1.  
2. Ces deux modèles ont été décrits plus en détail à la section 2.3 du chapitre 2.  
3. *Cat's Claw* est disponible en ligne à l'adresse <http://www.cs.uky.edu/~raphael/linguistics/claw.html> (valable au 11 juin 2013).  
4. <http://networkmorphology.as.uky.edu> (valable au 11 juin 2013).

d'héritage non monotone comme DATR (Evans & Gazdar, 1989 ; Corbett & Fraser, 1993 ; Andry et al., 1993 ; Barg, 1994) et ses extensions comme KATR (Finkel & Stump, 2002), une grande majorité des approches computationnelles relèvent de la Morphologie à États Finis (Beesley & Karttunen, 2003). Ces approches s'intéressent principalement, voire souvent exclusivement, aux moyens computationnels pour traiter différentes opérations morphologiques depuis la simple concaténation (affixation) jusqu'à la réduplication en passant par la morphologie gabaritique.

Roark & Sproat (2007, chap. 3) évaluent la pertinence des questions computationnelles pour la théorie morphologique, en partie pour le débat opposant les défenseurs des approches *Item et Arrangement (IA)* à ceux des approches *Item et Processus (IP)*. Ils arrivent à la conclusion que 1) d'un point de vue computationnel les approches IA et IP sont également expressives et toutes deux exprimables par des transducteurs à états finis, et 2), que les théories morphologiques qu'ils analysent, et en particulier celle de Stump (2001), sont réductibles à des transducteurs à états finis<sup>5</sup>.

Ce constat indique ainsi que quelque soit le cadre théorique<sup>6</sup> choisi, il est possible de décrire la morphologie et de générer l'intégralité des formes par des transducteurs à états finis<sup>7</sup>. En vertu du principe du rasoir d'Ockham, ils proposent de couper court au débat théorique et de choisir la formalisation la plus simple rendant compte des données, en l'occurrence selon eux la Morphologie à États Finis.

Mais ce constat, certes pertinent, ne décrit en fait que le minimum de ce qu'on est en droit d'attendre d'un modèle morphologique, à savoir qu'il rende compte de façon exacte des réalités observables de la langue. Au delà de ce minimum, d'autres questions se posent qui différencient entre elles les approches et leur donnent de la valeur supplémentaire, potentiellement dans des domaines différents selon les approches choisies. L'approche que nous avons choisie pour le développement de notre modèle cherche, au delà de la génération correcte des formes, à rendre compte des régularités observables à l'intérieur du paradigme. La façon dont les règles de réalisation sont effectivement implémentées y est en fin de compte assez secondaire.

Les constats de Karttunen et de Roark & Sproat rejoignent cependant ce que nous affirmions plus haut : compte tenu de la possibilité de concevoir des descriptions formelles

---

5. Voir aussi avant eux Karttunen (2003) ainsi que Malouf (2005).

6. Notons qu'ils n'étudient évidemment pas l'intégralité des cadres théoriques existants. Parmi les approches contemporaines majeures, les approches apparentées à la Morphologie Distribuée ne sont en particulier pas évoquées par les auteurs.

7. Exception doit être faite ici des cas de langues dont une partie de la morphologie emploie la réduplication comme procédé morphologique, ce dernier ne pouvant pas être couvert de façon générique par des systèmes employant tels quels les transducteurs à état finis.

de la morphologie qui rendent compte de façon exacte de l'intégralité du système flexionnel d'une langue, on est en droit d'être sceptique quant aux approches qui ne satisfont pas au moins cette exigence de base. Par ailleurs nous considérons que les approches computationnelles constituent une base solide pour la conception de modèles théoriques cherchant à exprimer des réalités complémentaires, tel que, par exemple, la caractérisation typologique d'un système particulier. En cela nous souscrivons donc à ce qu'affirme Karttunen (2003), à savoir que la combinaison des approches computationnelles et théoriques apparaît alors comme la méthode la plus souhaitable.

### 5.1.3 De l'importance de la grande échelle

L'implémentation des analyses en morphologie flexionnelle permet de vérifier la cohérence de l'analyse proposée. Une implémentation à grande échelle permet de surcroît de garantir la quasi-exhaustivité de la description proposée. Elle permet notamment de certifier la cohérence globale d'une analyse par rapport à l'intégralité du système et à rendre compte du statut exceptionnel ou au contraire étendu d'un phénomène donné.

Un exemple de cela est l'analyse de la flexion des adjectifs du français par Bonami & Boyé (2005). Les auteurs y discutent la formation de la forme de liaison du masculin singulier FLMS. La FMLS est généralement identique à la forme du masculin singulier dès lors que ce dernier se termine par une consonne et qu'il est identique au féminin singulier dès lors que le masculin singulier se termine par une voyelle. Les auteurs illustrent ce comportement par les exemples indiqués au tableau 5.1.

Nous donnons également un exemple des trois formes du singulier de quelques adjectifs du français proposés par Bonami & Boyé au tableau 5.2.

Il est par ailleurs connu que pour certains adjectifs comme GRAND ou GROS, la FMLS peut différer autant du M.SG que du F.SG. La FMLS de GROS sera ainsi [gʁoz] et non pas [gros] comme le F.SG, tel qu'attendu pour un adjectif dont le M.SG [gʁo] se termine par une voyelle. Il en va de même pour la FMLS [gʁãt] de GRAND par rapport à sa forme [gʁãd] au F.SG et [gʁã] au M.SG. La FMLS présente pour les adjectifs de ce type une alternance de consonne voisée/non voisée par rapport à la forme du F.SG. L'explication de cette situation occupe une place centrale dans la littérature en phonologie théorique. Bonami & Boyé (2005, p. 81) montrent cependant que ce type d'alternance ne concerne en fin de compte qu'un nombre très petit de lexèmes : quatre adjectifs à alternance s/z (BAS, DOUX, FAUX et GROS) et trois adjectifs à alternance t/d (GRAND, PROFOND et SECOND) sur l'inventaire

	M.SG	FMLS	F.SG
SEC	sec rendez-vous [sɛkʁãdevu]	sec entretien [sɛkãtɛtjẽ]	sèche entrevue [sɛʃãtɛvvy]
COURT	court rendez-vous [kuʁãdevu]	court entretien [kuʁãtɛtjẽ]	courte entrevue [kuʁtãtɛvvy]
VIF	vif sentiment [vifsãtimã]	vif émoi [vifemwa]	vive émotion [vivemosjõ]
VIEUX	vieux bateau [vjøbato]	vieil avion [vjɛjavjõ]	vieille auto [vjɛjoto]
NOUVEAU	nouveau bateau [nuvobato]	nouvel avion [nuvelavjõ]	nouvelle auto [nuveloto]
BEAU	beau bateau [bobato]	bel avion [belavjõ]	belle auto [bøloto]
PETIT	petit bateau [pɛtibato]	petit avion [pɛtitavjõ]	petite auto [pɛtitoto]
GRAND	grand bateau [gʁãbato]	grand avion [gʁãtavjõ]	grande auto [gʁãdoto]
GROS	gros bateau [gʁobato]	gros avion [gʁozavjõ]	grosse auto [gʁosoto]

TABLE 5.1 : Exemples de FMLS de quelques adjectifs du français (Bonami & Boyé, 2005)

	M.SG	F.SG	FLMS
JOLI	ʒoli	ʒoli	ʒoli
NET	nɛt	nɛt	nɛt
PETIT	pɛti	pɛtit	pɛtit
COURT	kuʁ	kuʁt	kuʁt
NOUVEAU	nuvo	nuvel	nuvel
GRAND	gʁã	gʁãd	gʁãt
CHAUD	ʃo	ʃod	–

TABLE 5.2 : Sous-paradigme du singulier de quelques adjectifs du français (Bonami & Boyé, 2005)

complet des adjectifs répertoriés par le TLFi<sup>8</sup>. Il apparaît ainsi qu'il n'est pas nécessaire de postuler une règle spécifique dans la grammaire, dans la mesure où il est bien plus économique de spécifier ces formes spécifiques dans le lexique comme des idiosyncrasies de quelques rares entrées lexicales. En tant que phénomène marginal, ce phénomène ne mérite pas d'être décrit comme une propriété du système flexionnel des adjectifs au même titre que ce qui conditionne pour la majeure partie du lexique l'alternance entre FMLS identique au M.SG, respectivement au F.SG, selon la structure phonologique du M.SG. Ladite marginalité du phénomène n'apparaît cependant que dès lors que l'on confronte le nombre de lexèmes concernés à un lexique de taille réaliste.

De même, l'exemple de l'analyse du maltais que nous donnerons à la section 5.3 montrera comment seule la prise en compte de la quasi-exhaustivité des données d'un sous-système flexionnel du maltais a pu permettre de mettre en évidence des points que des analyses théoriques partielles préliminaires n'avaient pas prises en compte. Ces faits montrent ainsi clairement l'avantage à approcher la morphologie flexionnelle en prenant en compte des données quasi-exhaustives sur le phénomène étudié.

## 5.2 Implémenter $\mathcal{PARSL}$

Afin de permettre l'élaboration de descriptions morphologiques dans notre formalisme qui satisfassent la contrainte de l'implémentation à grande échelle, nous avons choisi le formalisme lexical Alexina (Sagot, 2010) comme fondement de l'implémentation de  $\mathcal{PARSL}$ . Ce choix présente trois avantages. 1) Le formalisme Alexina est un formalisme développé pour le traitement automatique des langues. Il a déjà fait ses preuves en tant que formalisme lexical, un certain nombre de lexiques à grande et moyenne échelle ayant déjà été développés dans ce formalisme. Le plus important est le *Lefff* (Sagot et al., 2006; Sagot, 2010), lexique du français (écrit) comportant près de 120 000 lemmes pour environ 550 000 formes et 460 000 formes distinctes. D'autres lexiques Alexina sont répertoriés au tableau 5.3. 2) Tous les lexiques Alexina sont des lexiques libres. Les implémentations d'analyses en  $\mathcal{PARSL}$  ayant donné lieu à de nouveaux lexiques Alexina $\mathcal{PARSL}$ <sup>9</sup> sont ainsi également librement disponibles<sup>10</sup>. Ce fait garantit la vérifiabilité des analyses autant que la possibilité de réemployer ces lexiques pour des travaux ultérieurs

8. <http://atilf.atilf.fr> (valable au 11 juin 2013).

9. Cases grisées dans le tableau 5.3.

10. Les analyses sont disponibles en ligne sur notre page personnelle. L'intégralité des lexiques Alexina est par ailleurs disponible sur la page du projet Alexina : <http://alexina.gforge.inria.fr> (valable au 11 juin 2013).

en traitement automatique des langues par l'intégralité de la communauté. 3) Alexina comporte explicitement une distinction articulée entre lexique flexionnel et grammaire flexionnelle similaire à celle définie dans  $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$ . L'implémentation de cette distinction peut ainsi se faire très naturellement en Alexina $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$ .

LEXIQUE	LANGUE	LEMES	LEXÈMES	FORMES	F. DIST.	RÉFÉRENCE
<i>Lefff</i> <sup>11</sup>	français	120 000	125 000	550 000	460 000	(Sagot, 2010 ; Walther & Sagot, 2011)
<i>Leffe</i>	espagnol	180 000	180 000	1 500 000	700 000	(Molinero et al., 2009)
<i>Leffga</i>	galicien	70 000	70 000	750 000	500 000	(Molinero et al., 2009)
<i>Leffla</i>	latin	2 200	2 200	115 000	96 000	non publié
EnLex	anglais	350 000	350 000	580 000	510 000	(Sagot, 2010)
PolLex	polonais	240 000	240 000	1 400 000	360 000	(Sagot, 2007)
SkLex	slovaque	50 000	50 000	470 000	250 000	(Sagot, 2005)
PerLex	persan	30 000	30 000	550 000	460 000	(Sagot et al., 2011 ; Sagot & Walther, 2010)
KurLex	kurde kurmanji	22 000	22 000	410 000	240 000	(Walther et al., 2010)
SoraLex	kurde sorani	520	520	30 000	25 000	(Walther & Sagot, 2010)
MaltLex	maltais	560	560	9 000	7 200	(Camilleri & Walther, 2012)

TABLE 5.3 : Les lexiques Alexina

Dans cette section nous décrivons le développement d'Alexina $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$ , un système d'implémentation pour les analyses formelles dans le modèle  $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$  qui repose sur une version modifiée du formalisme lexical Alexina (Sagot, 2010).

### 5.2.1 Le formalisme Alexina originel

Le formalisme Alexina est un formalisme lexical développé depuis 2003 (Clément et al., 2004 ; Sagot, 2005 ; Sagot et al., 2006 ; Sagot, 2006, 2007, 2010). Il comporte à la fois un niveau morphologique et un niveau syntaxique. Nous ne nous intéresserons ici toutefois qu'au niveau morphologique, effectivement affecté par l'adaptation d'Alexina au modèle

11. La version du *Lefff* en Alexina $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$  employant les notions de  $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$  est celle pulée dans (Sagot & Walther, 2011) et (Walther & Sagot, 2011). La version originale et officielle du *Lefff* telle que publiée dans (Sagot, 2010) repose cependant quant à elle sur la version antérieure d'Alexina.

PARSLI. Notons aussi qu'Alexina prévoit également un encodage des opérations de dérivation possibles. Nous ne détaillons cependant pas plus ce mécanisme, la modification d'Alexina pour prendre en compte les notions de PARSLI ne concernant que la flexion.

Alexina est un formalisme lexical développé pour le traitement automatique des langues (TAL). Elle s'accompagne également d'une panoplie d'outils pour l'enrichissement automatique de lexiques (Sagot, 2007).

Sagot (2006, p. 84) défend le point de vue selon lequel, en TAL, une description morphologique doit *a priori* satisfaire les critères de PERTINENCE LINGUISTIQUE, de RÉVERSIBILITÉ, de COMPACTITÉ et de LISIBILITÉ. Alexina a par conséquent été développée dans le but de satisfaire cette liste de critères.

Le critère de pertinence linguistique avait déjà poussé Sagot à proposer des analyses surfaciques de la grammaire en général et de la morphologie en particulier. Il cherche ainsi au maximum à éviter l'emploi de concepts abstraits tels que parfois employés dans d'autres approches computationnelles à la morphologie, comme le « 0 » en *Morphologie à deux niveaux* (Koskenniemi, 1984). Dans la nouvelle version du formalisme Alexina que nous avons développée, nous avons encore renforcé cette exigence théorique.

Le critère de réversibilité concerne principalement l'utilisation de descriptions morphologiques non seulement pour fléchir des entrées lexicales mais également pour *défléchir* de façon ambiguë des formes fléchies, par exemple pour construire de nouvelles entrées lexicales candidates. Pour le choix du formalisme lexical sous-jacent à l'implémentation de notre modèle morphologique, nous soulignons ici simplement que cette propriété d'Alexina facilite l'acquisition de lexiques à grande échelle et par conséquent l'application des analyses implémentées à une liste quasi-exhaustive d'entrées lexicales de manière à garantir la cohérence globale du système morphologique décrit.

Le critère de compacité prévoit que, toutes choses égales par ailleurs, une description morphologique factorise maximale l'information commune à plusieurs classes flexionnelles et évacue les redondances inutiles. Cette factorisation ne doit cependant pas se faire au détriment de la lisibilité de la description par des descripteurs humains. Une description morphologique implémentée doit par conséquent être formulée et représentée de sorte à rester compréhensible pour d'autres linguistes qui voudraient l'amender, la corriger ou la compléter ultérieurement. Alexina avait ainsi dès le départ vocation à servir de formalisme pour la description morphologique en général (et pas seulement pour le TAL). Néanmoins, le formalisme Alexina ne va de ce point de vue pas assez loin, notamment dans la définition de catégories et d'étiquettes compatibles avec celles employées dans les approches théoriques. En définissant Alexina<sub>PARSLI</sub>, nous avons tenté de remédier à cela et

de ne garder dans le nouveau formalisme lexical que des catégories ayant par ailleurs une définition théorique précise au sein de  $\mathcal{PARSLI}$ .

### 5.2.1.1 Représentation de la morphologie et du lexique dans Alexina

La représentation de la morphologie dans Alexina repose explicitement sur une approche paradigmatique de la morphologie. Chaque entrée lexicale spécifiée est ainsi associée à une classe flexionnelle. Une classe flexionnelle, spécifiée par la grammaire, comporte les règles de réalisation permettant de construire l'intégralité de son paradigme.

Le formalisme lexical Alexina est un formalisme lexical qui s'appuie sur deux niveaux de représentation : un NIVEAU INTENSIONNEL d'une part, et un NIVEAU EXTENSIONNEL de l'autre. Le niveau intensionnel comporte les entrées lexicales avec leur assignation de classes flexionnelles.. Nous donnons ci-dessous cinq exemples d'entrées morphologiques verbales intensionnelles extraites du *Lefff* (version standard)<sup>12</sup>. Chaque entrée est définie par sa forme de citation (respectivement « *accoutumer* », « *acquérir* » et « *admettre* ») et sa CLASSE FLEXIONNELLE associée. Les classes flexionnelles en question sont `v-er` pour les verbes du premier groupe réguliers, `v-ir2` pour les verbes du deuxième groupe en *-ir*, `v-ir3` pour les verbes du troisième groupe en *-ir* et `v55` pour une des tables de verbes irréguliers du troisième groupe. À ces classes s'ajoutent pour certains verbes des VARIANTES permettant de définir des règles spécifiques pour un sous-ensemble d'entrées lexicales appartenant à une classe flexionnelle donnée. Dans l'exemple de la figure 5.1 les variables `dbl` et `std` s'appliquent respectivement aux entrées des verbes du premier groupe qui redoublent la consonne finale de leur radical (cf. *appeler* par opposition à *appelle*) et aux verbes du premier groupe standard qui ne sont pas soumis à ce redoublement consonnantique (cf. *accoutumer* et *accoutume*).

```
1 accoutumer v-er:std
2 appeler v-er:dbl
3 enrichir v-ir2
4 dormir v-ir3
5 admettre v55
```

FIGURE 5.1 : Entrées lexicales verbales intensionnelles extraites du *Lefff*

Ces entrées peuvent être compilées à l'aide de la grammaire afin de produire le lexique extensionnel comportant l'intégralité des formes fléchies des entrées du

12. Pour plus de lisibilité, nous avons supprimé les informations syntaxiques spécifiées dans le lexique intensionnel.



lexique intensionnel<sup>13</sup>. Nous donnons à la figure 5.2 trois exemples partiels d'entrées extensionnelles compilées à partir des entrées intensionnelles à la figure 5.1<sup>14</sup>. Notons que les étiquettes morphosyntaxiques comme J3s employées dans le *Lefff* original s'inspirent du standard MULTTEXT (Ide & Véronis, 1994) : en l'occurrence, l'étiquette J désigne le passé simple, P le présent indicatif, G le participe présent, s le singulier et 1 et 3 respectivement la première et la troisième personne. La notation appeler\_\_\_\_\_1 permet de dissocier le lexème « *appeler* » au sens de '*interpeller quelqu'un par son nom*' du lexème « *appeler* » homonyme noté appeler\_\_\_\_\_2 et signifiant '*téléphoner à quelqu'un*'.

```

1 accoutuma v accoutumer_____1 J3s
2 accoutume v accoutumer_____1 PS13s
3 accoutumant v accoutumer_____1 G
4
5 appela v appeler_____1 J3s
6 appelle v appeler_____1 PS13s
7 appelant v appeler_____1 G
8
9 enrichit v enrichir_____1 J3s
10 enrichit v enrichir_____1 P3s
11 enrichissant v enrichir_____1 G
12
13 dormit v dormir_____1 J3s
14 dort v dormir_____1 P3s
15 dormant v dormir_____1 G
16
17 admit v admettre_____1 J3
18 admet v admettre_____1 P3
19 admettant v admettre_____1 G

```

FIGURE 5.2 : Entrées lexicales verbales extensionnelles produites à partir du *Lefff*

De même, les entrées intensionnelles et extensionnelles adjectivales sont représentées comme à la figure 5.3.

```

1 bordeaux adj-1
2 rapide adj-2
3 grand adj-4

```

FIGURE 5.3 : Entrées lexicales adjectivales intensionnelles extraites du *Lefff*

13. C'est également le lexique extensionnel qui sert d'entrée aux outils de traitement automatique des langues comme les analyseurs syntaxiques et certains étiqueteurs morphosyntaxiques.

14. Pour plus de lisibilité, nous avons supprimé les informations syntaxiques également spécifiées dans le lexique extensionnel.

Pour les adjectifs, nous donnons à la figure 5.4 l'intégralité des formes fléchies distinctes générées par le *Lefff* pour les trois entrées intensionnelles ci-dessus. Les étiquettes *m* et *f* désignent respectivement les valeurs de traits masculin et féminin.

```

1 bordeaux adj bordeaux_____1
2
3 rapide adj rapide_____1 s
4 rapides adj rapide_____1 p
5
6 grand adj grand_____1 ms
7 grand' adj grand_____1 s
8 grande adj grand_____1 fs
9 grandes adj grand_____1 fp
10 grands adj grand_____1 mp

```

FIGURE 5.4 : Entrées lexicales adjectivales extensionnelles produites à partir du *Lefff*

À côté du lexique extensionnel, Alexina permet enfin de spécifier manuellement les formes irrégulières qui ne sont pas générées par la grammaire. Ces formes se rajoutent aux formes générées de façon régulière par la grammaire à partir du lexique intensionnel.

La grammaire comporte quant à elle deux parties essentielles : une partie phonographémique et une partie proprement morphologique. Les règles phonographémiques simulent la morphologie pour des lexiques de formes écrites. Cette composante est néanmoins également employable pour représenter des règles morphologiques au sens strict. Alexina permet par ailleurs de définir des classes de graphèmes (ou de phonèmes) auxquels des règles peuvent s'appliquer. Nous donnons à la figure 5.5 un exemple pour deux des classes des voyelles et des consonnes (graphémiques) du *Lefff*.

```

1 <letterclass name="voyelle" letters="a à â ä e é è ê ë i î ï o ô ö u û
   ü ù y ÿ"/>
2 <letterclass name="consonne" letters="b c ch d f gu g h j k gl l m gn
   n nj p qu q tr br vr cl cr gr r s t v w x z"/>

```

FIGURE 5.5 : Classes de graphèmes/phonèmes en Alexina (données du *Lefff* original)

Ces règles peuvent ensuite servir dans les règles morphographémiques ou morphologiques comme celles du *Lefff* indiquées à la figure 5.6. Elles indiquent respectivement que le graphème <g> est remplacé par la séquence <ge> lorsqu'il est suivi d'un suffixe à initiale vocalique de type <a>, <o> ou <u><sup>15</sup>. De même <c> devient <ç> dans ce même contexte, tandis que <qu> devient <c> devant consonne.

15. Dans les règles morphographémiques/morphologiques, Alexina marque la frontière de morphe par

```

1 %le:]" target="quière[:voyelle:]" />
2 % <sandhi source="quér-r" target="quer_r" />
3 % <sandhi source="quér_-[:consonne:]" target="quier_[:consonne:]" />
4
5 <sandhi source="g_[:aou:]" target="ge_[:aou:]" />
6 <sandhi source="c_[:aou:]" target="ç_[:aou:]" />
7 <sandhi source="qu_[:consonne:]" target="c_[:consonne:]" />

```

FIGURE 5.6 : Règles morphographémiques/morphologiques en Alexina (données du *Lefff* original)

La partie proprement morphologique comporte des classes flexionnelles construites comme suit. Dans l'exemple des entrées adjectivales donné à la figure 5.3, nous pouvons constater que le nombre de formes adjectivales distinctes générées par la grammaire diffère d'une classe adjectivale à une autre. La classe *adj-1* est invariable, tandis que la classe *adj-4* permet de générer quatre formes distinctes. Nous donnons à la figure 5.7 les exemples de classes flexionnelles des adjectifs indiqués plus haut.

La table *adj-1* est celle des adjectifs invariables. Elle se caractérise par le fait qu'elle ne modifie aucunement la forme de citation, employée comme forme de base. Ce fait est représenté en Alexina original par une opération de suffixation vide (*suffix=""*). Par ailleurs, aucun trait particulier n'est exprimé par la forme générée, ce qui est obtenu par une absence d'étiquette (*tag=""*) dans l'unique règle de réalisation de forme indiquée dans la table associée aux adjectifs invariables dans la description morphologique.

```

1 <table name="adj-1" rads="..*">
2 <form suffix="" tag="" />
3 </table>

```

FIGURE 5.7 : La classe flexionnelle *adj-1* dans la description morphologique du *Lefff* original

Les autres classes flexionnelles peuvent quant à elles comporter des règles de réalisation de type préfixation ou suffixation et associer des étiquettes morphosyntaxiques à chaque forme ainsi générée.

Elles peuvent par ailleurs également emprunter les règles d'autres classes, y compris de classes fléchissant des catégories syntaxiques différentes. Ainsi, la classe *adj-4* emprunte ses règles à la classe *nc-4* des noms masculin ayant un correspondant féminin en *-e* comme

---

le tiret bas.

« *doctorant/doctorante* <sup>16</sup> ». Nous indiquons cette dernière à la figure 5.8 <sup>17</sup> :

```

1 <table name="nc-4" rads="..*[^sxzce]">
2   <form suffix="" tag="ms" />
3   <form suffix="e" tag="fs" />
4   <form suffix="s" tag="mp" />
5   <form suffix="es" tag="fp" />

```

FIGURE 5.8 : La classe flexionnelle `nc-4` dans la description morphologique du *Lefff* original

La table `adj-4` (figure 5.9) emprunte ainsi les règles de la classe `nc-4` à l’aide de la règle de renvoi `<like name="nc-4"/>` <sup>18</sup>.

```

1 <table name="adj-4" rads="..*[^esc]">
2   <like name="nc-4"/>
3 </table>

```

FIGURE 5.9 : La classe flexionnelle `adj-4` dans la description morphologique du *Lefff* original

Les règles de renvoi dans le formalisme Alexina d’origine n’existent par ailleurs pas que pour des renvois entre tables, mais également pour des renvois entre règles d’une même table. Un exemple de cela est le cas de la règle de renvoi entre le génitif singulier et l’accusatif singulier dans la flexion nominale des noms masculins animés qui ne se terminent pas en « *a* » au nominatif singulier du polonais, illustré à la figure 5.10.

Enfin, Alexina permet par le mécanisme des variantes évoqué plus haut de spécifier une sous-classe des lexèmes appartenant à une classe flexionnelle donnée pour laquelle certaines règles s’appliquent, les autres entrées en étant exclues. Un exemple de cela

16. Notons que cet exemple présente les paradigmes des noms comme « *doctorant* » comme des paradigmes à quatre cases avec une variation en genre. Cet exemple a été choisi ici à titre purement illustratif. Nous n’affirmons ainsi aucunement qu’il s’agit là de la représentation la plus appropriée de la flexion nominale du français.

17. La spécification `rads="..*[^sxzce]"` indique que les entrées lexicales se fléchissant selon cette classe doivent satisfaire les contraintes suivantes : ils doivent ne pas se terminer en « *s* », « *x* », « *z* », « *c* » ou « *e* ». Alexina possède également un autre mécanisme de sélection d’entrées compatibles avec une classe flexionnelle particulière. La mention `except=".e"` dans `<table name="nc-2m" rads="..*[^sxz]" except=".e">` exclut des candidats possibles à une flexion par la classe `nc-2m` les noms de deux caractères dont le second est « *e* ».

18. Le mécanisme d’exclusion `except` peut également être utilisé pour restreindre une règle de renvoi en excluant des parties de classes flexionnelles n’étant pas empruntées par ladite règle, cf. `<like name="v-ir2" except="PJ12s"/>` dans la table empruntant l’intégralité des règles de la table `v-ir2` à l’exclusion de la première et deuxième personne du singulier de l’indicatif du présent et du passé simple.

```
1 <form like="sg:gen" tag="sg:acc" />
```

FIGURE 5.10 : Exemple d'une règle de renvoi intra-classe en Alexina extraite de PolLex

est le redoublement de la consonne finale du radical de certains verbes du premier groupe dans certaines cases du paradigme comme le 1SG.PRS.IND et le 3SG.PRS.IND. Cette caractéristique est représentée comme suit.

Les entrées lexicales verbales du premier groupe peuvent être spécifiées dans le lexique intensionnel par une variante `dbl` pour les verbes à redoublement de consonne ou `std` pour les autres. Les règles de réalisation peuvent également spécifier une variante : dans l'exemple que nous donnons ci-dessous, ce sont les variantes `var="std"` et `var="dbl"`. Les entrées lexicales ne se fléchiront alors que par les règles qui ne comportent pas de variante incompatible avec la leur. Les règles comportant des variantes exclusives sont regroupées à l'intérieur de groupes de règles dont une et une seule règle peut à chaque fois qu'être appliquée à une entrée lexicale donnée<sup>19</sup>. Cette notation est illustrée à la figure 5.11.

```
1 <alt>
2 <form suffix="2e" tag="PS13s" var="dbl" />
3 <form suffix="e" tag="PS13s" var="std" />
4 </alt>
```

FIGURE 5.11 : Règles de réalisations incompatibles et variantes en Alexina (*Lefff* original)

### 5.2.2 Le formalisme Alexina adapté à $\mathcal{PARSL}$

Pour établir un système d'implémentation pour  $\mathcal{PARSL}$ , nous sommes ainsi partis du formalisme Alexina d'origine et nous l'avons adapté afin qu'il prenne en compte les notions propres à  $\mathcal{PARSL}$ . Dans l'adaptation d'Alexina au modèle morphologique  $\mathcal{PARSL}$ , nous avons tenté de maintenir, voire d'améliorer, au maximum les exigences pour les descriptions morphologiques posées par Sagot (2006) : réversibilité, compacité et lisibilité. Le critère de la réversibilité n'ayant entraîné dans le développement du nouveau formalisme que des difficultés d'implémentation, certes complexes, mais non essentielles pour l'objet de cette thèse, nous ne développerons plus en détail dans la suite de cette section que le cas des exigences de compacité et de lisibilité.

Le nouveau formalisme Alexina $\mathcal{PARSL}$  comporte les innovations suivantes par rapport au formalisme d'origine.

19. Ces groupes exclusifs sont indiqués par les balises `<alt> ... </alt>` dans une description Alexina.

1. Le lexique intensionnel d'Alexina<sub>PARSLI</sub> représente le lexique flexionnel dans  $\mathcal{PARSLI}$ ; la grammaire morphologique d'Alexina<sub>PARSLI</sub> correspond à la grammaire flexionnelle dans  $\mathcal{PARSLI}$ .
2. Grâce à l'introduction d'un mécanisme de définition d'opérations sur les formes, Alexina<sub>PARSLI</sub> permet désormais de représenter directement des opérations morphologiques non concaténatives.
3. Alexina<sub>PARSLI</sub> repose sur un mécanisme explicite de définition des traits morphosyntaxiques permettant de calculer l'intégralité des cases des paradigmes flexionnels.
4. Alexina<sub>PARSLI</sub> implémente le système des niveaux morphologiques typés en niveaux de type radical, thématique et exposant.
5. La spécification des classes flexionnelles est devenue optionnelle, seules les zones sont obligatoires. Toutefois, pour des raisons de rétrocompatibilité, lorsqu'aucune zone n'est spécifiée dans la grammaire, mais qu'une table est donnée, cette dernière est interprétée comme constituée d'une zone unique du même nom.
6. Alexina<sub>PARSLI</sub> permet également de définir des espaces partitionnants.
7. Les entrées lexicales sont associées à des schèmes dans le lexique intensionnel.
8. Alexina<sub>PARSLI</sub> permet de représenter directement dans le lexique intensionnel les cas de supplétion de radicaux.
9. Alexina<sub>PARSLI</sub> permet de rendre compte des cas de défektivité dans le lexique.
10. Les règles de réalisation sont appliquées si les traits qu'elles expriment sont unifiables avec ceux déjà exprimés par une forme donnée.

Les notions de variante, de classes de graphèmes/phonèmes et de règles morphographémiques/morphologiques ont été directement importées du formalisme original. L'application de la notion de variante a toutefois été généralisée : dans Alexina<sub>PARSLI</sub>, des variantes peuvent être spécifiées pour des règles de réalisation, des schèmes et des zones réalisationnelles à l'intérieur des schèmes.

### 5.2.2.1 Représentation des radicaux multiples dans Alexina<sub>PARSLI</sub>

Dans le formalisme Alexina d'origine, chaque entrée lexicale est *a priori* prévue pour n'avoir qu'un seul radical. Pour rendre compte de cas d'allomorphie radicale, il est nécessaire de contourner cette contrainte en détournant de leur utilité première des mécanismes disponibles. Trois types de solutions peuvent être envisagées, aucune d'entre elles n'étant linguistiquement satisfaisante.

1. Il est possible de lister les formes allomorphiques dans le fichier contenant les formes ajoutées manuellement et de ne pas les générer avec la grammaire. Cette solution requiert d'associer par ailleurs les entrées lexicales concernées par ce type d'allomorphie à des classes flexionnelles défectives pour les cases comportant de l'allomorphie.
2. On peut associer les entrées supplétives à des classes flexionnelles dans lesquelles le radical de l'entrée en question est considéré comme un segment vide, ou un segment de taille *submorphique*, et les formes complètes sont modélisées comme des suffixes s'accolant à ce radical vide (ou submorphique). La forme entière est alors traitée comme la désinence correspondant à une case donnée.
3. On peut détourner l'emploi des règles morphographémiques/morphonologiques et en faire des transformateurs de radicaux. Ainsi, quand le radical par défaut est combiné avec un affixe habituellement adjoint à un radical allomorphique, la règle morphographémique/morphonologique simule l'allomorphie en transformant la partie radicale. Cette solution d'appoint a notamment été employée pour simuler l'allomorphie radicale en persan et en kurde (Sagot et al., 2011 ; Sagot & Walther, 2010 ; Walther & Sagot, 2010 ; Walther et al., 2010) avant la mise en place d'Alexina<sub>PARSLI</sub>.

Dans le nouveau formalisme Alexina<sub>PARSLI</sub>, l'allomorphie peut être prise en compte par deux mécanismes distincts, selon qu'il s'agit d'une allomorphie régulière comme dans les langues iraniennes ou de supplétion ponctuelle, comme pour la supplétion de radicaux du verbe *aller* en français pour lequel les descriptions indiquent jusqu'à quatre radicaux différents ; *all-*, *v-*, *i-* et *aill-*. Nous donnons aux figures 5.12 et 5.13 (pages 294 et 295) la représentation de ces deux alternances de radicaux telles qu'implémentées en Alexina<sub>PARSLI</sub>.

L'allomorphie régulière est donnée par la définition de règles de réalisation propres au niveau radical dans la grammaire. La figure 5.12 donne un exemple de ces règles pour l'allomorphie régulière en maltais. Un extrait des données du maltais concernées est indiqué au tableau 5.4 page 293. Elles montrent que les verbes maltais affichent une alternance de radicaux régulière<sup>20</sup>. Parmi les paradigmes existant en maltais, il peut y avoir de l'allomorphie radicale distinguant jusqu'à six radicaux différents (RAD1 à RAD6). Dans les cas des deux paradigmes indiqués au tableau 5.4 l'alternance correspond à quatre radicaux différents indiqués par les couleurs des cases, deux des six radicaux potentiellement différenciables étant syncrétiques pour les paradigmes concernés : le RAD1 et le RAD3 d'une part et le RAD5 et le RAD6 de l'autre.

20. Décrite plus en détail à la section 5.3.

		RASS « bloquer »	MESS « toucher »
RAD2	PFV 1SG	<i>rasséjt</i>	<i>messéjt</i>
	PFV 2SG	<i>rasséjt</i>	<i>messéjt</i>
	PFV 1PL	<i>rasséj.na</i>	<i>messéj.na</i>
	PFV 2PL	<i>rasséj.tu</i>	<i>messéj.tu</i>
RAD1	PFV 3.M.S	<i>ráss</i>	<i>méss</i>
RAD3	PFV 3.F.S	<i>rás.set</i>	<i>més.set</i>
RAD4	PFV 3PL	<i>ras.sé:w</i>	<i>mes.sé:w</i>
RAD5	IPFV 1SG	<i>nróss</i>	<i>nmíss</i>
	IPFV 2SG	<i>tróss</i>	<i>tmíss</i>
	IPFV 3.M.S	<i>jróss</i>	<i>jmíss</i>
	IPFV 3.F.S	<i>tróss</i>	<i>tmíss</i>
RAD6	IPFV 1PL	<i>nrós.su</i>	<i>nmís.su</i>
	IPFV 2PL	<i>trós.su</i>	<i>tmís.su</i>
	IPFV 3PL	<i>jrós.su</i>	<i>jmís.su</i>

TABLE 5.4 : Paradigmes des verbes RASS et MESS en maltais

L'alternance de radicaux est alors rendue dans la grammaire par des opérations sur les radicaux. En maltais, la forme de citation (PFV.3.S.M) est identique à l'un des radicaux : aucune marque d'exponence n'est apparente sur cette forme. Dans l'implémentation que nous avons faite de la flexion verbale maltaise, nous avons ainsi appelé RAD1 ce radical<sup>21</sup>. Les autres radicaux sont obtenus par des opérations sur RAD1. Ainsi, le RAD2 s'obtient par l'ajout à droite de la séquence « *ej* ». Cette opération est rendue par la règle prenant en entrée le RAD1 (source="S1") appliquant l'opération `append="ej"` et générant ainsi le RAD2 (name="S2"). De même, le RAD5 (name="S5") est obtenu à partir du RAD1 par l'opération `operation="changeV1(i)"` appliquant un changement vocalique e→i pour les verbes ayant la voyelle « *e* » dans leur forme de citation<sup>22</sup> comme MESS « *toucher* » et par l'opération `operation="changeV1(o)"` appliquant un changement vocalique a→o pour les verbes ayant la voyelle « *a* » dans leur forme de citation comme RASS « *bloquer* ». Les syncrétismes entre deux radicaux sont indiqués par des opérations vides (`operation=""`). Deux exemples de syncrétismes sont ainsi le cas de la formation du

21. Alexina<sup>ᠮᠠᠯᠠᠰᠢ</sup> définit en fait un RADICAL 0 S0 qui par défaut est identique au S1. S0 sert explicitement dans les cas où le S1 ne permet pas de construire les autres radicaux ou dans les cas de surabondance de radical au niveau RAD1, cf. le cas de certains paradigmes maltais décrits à la section 5.3.

22. Cf. la contrainte sur les radicaux exprimée par `rads="[:C:]e[:C:][:C:]` indiquant une structure d'une consonne suivie de « *e* » suivi de deux autres consonnes.



RAD3 (S3) à partir du RAD1 (S1) et le cas de la formation du RAD6 (S6) à partir du RAD5 (S5)<sup>23</sup>.

```

1
2 <table name="CVCC" rads="[:C:][:V:][:C:][:C:]">
3   <item name="S1"/>
4
5   <item name="S2" source="S1" append="ej"/>
6
7   <item name="S3" source="S1" operation="" />
8
9   <item name="S4" source="S1" append="e"/>
10
11  <item name="S5" source="S1" operation="changeV1(o)" rads="[:C:]a[:C:]
12  :][:C:]" />
13  <item name="S5" source="S1" operation="changeV1(i)" rads="[:C:]e[:C:]
14  :][:C:]" />
15  <item name="S5" source="S1" operation="changeV1(i)" rads="[:C:]i[:C:]
16  :][:C:]" />
17
18  <item name="S6" source="S5" operation="" />
19
20 </table>

```

FIGURE 5.12 : Allomorphie radicale régulière en Alexina $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$  (données de MaltLex)

L'autre façon de rendre compte des radicaux multiples dans Alexina $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$  est de les lister dans le lexique. Ce procédé est conçu pour rendre compte des cas de supplétion de radical où la création d'un radical n'est pas prise en compte par la grammaire mais constitue une irrégularité lexicale extrême de l'entrée lexicale en question. Nous utilisons ce procédé pour rendre compte des radicaux supplétifs du verbe ALLER en français. Dans une implémentation du système verbal du français admettant l'existence d'un espace thématique à douze cases comme celui proposé par Bonami & Boyé (2003), nous rendons chaque case de l'espace thématique (c'est-à-dire chaque radical potentiellement différent pour un verbe donné) par un emplacement indiqué après le symbole / dans le lexique. Ces radicaux sont ordonnés et séparés les uns des autres par une virgule. Dans le cas de ALLER indiqué par l'entrée lexicale donnée dans la figure 5.13 page 295, les radicaux supplétifs spécifiés par le lexique sont ainsi le radical 2 « *va* », le radical 7 « *aill* » et le radical 10 « *i* ». Les autres radicaux sont déduits de façon régulière du radical « *all* » de la forme de citation ou par synchrétisme d'avec l'un des radicaux listés. Par exemple le radical 3 est déduit par synchrétisme avec le radical 2 et ainsi de suite.

23. Cette façon de construire un radical à partir d'un autre peut être comparée aux arbres de dépendance thématique au sens de Boyé (2000).

```
1 aller v:23r/,va,,,,,aill,,,i
```

FIGURE 5.13 : Supplétion de radical en Alexina (données du *Lefff* modifié)

### 5.2.2.2 La définition des niveaux réalisationnels dans Alexina<sub>PARSLI</sub>

Une autre nouveauté dans Alexina<sub>PARSLI</sub> est l'introduction des niveaux réalisationnels. Le formalisme Alexina d'origine ne comportait de règles de réalisation que pour le niveau d'exponence. Afin de pouvoir encoder l'allomorphie radicale régulière, nous avons introduit un niveau réalisationnel radical. Pour rendre compte des phénomènes non canoniques intervenant, pour certaines langues, au niveau thématique (selon le modèle  $\mathcal{PARSLI}$ ) nous avons également introduit un niveau thématique. Le résultat est l'existence d'une architecture à trois types de niveaux (`level`) dans Alexina<sub>PARSLI</sub>. Chaque niveau est typé (`type`) et ordonné par rapport aux autres (`level`). Ce dernier point a pour conséquence que la réalisation des formes se fait par l'application des règles de réalisations de chaque niveau dans l'ordre de 1 à  $n$ , où  $n$  est le dernier niveau défini dans une implémentation donnée pour une catégorie flexionnelle donnée. Nous donnons à la figure 5.14 page 296 un exemple des niveaux réalisationnels dans MaltLex comportant notre implémentation de la flexion verbale du maltais<sup>24</sup>. Les types de niveaux définis sont les niveaux `stem` pour le niveau radical, `theme` pour le niveau thématique et `exponent` pour le niveau exposant. Par ailleurs, comme dans le modèle  $\mathcal{PARSLI}$  à l'origine du nouveau formalisme Alexina<sub>PARSLI</sub>, seul le niveau radical est obligatoire. Si toutefois un mot est invariable, le niveau radical est par défaut interprété par le programme fléchissant les entrées lexicales comme unique et comportant uniquement une règle identifiant le radical à la forme de citation<sup>25</sup> indiquée dans le lexique.

24. Dans notre implémentation de la flexion verbale du maltais, chaque niveau est d'un type différent et il n'y a qu'un seul niveau par type (radical, thème, exposant). Si la description du système flexionnel d'une langue donnée le nécessitait, il pourrait tout à fait y avoir plusieurs niveaux de chaque type.

25. Formellement identique à ce radical.

```

1 <level type="stem" level="1">
2   .
3   <item name="S4" source="S3" append="" />
4   .
5   .
6   .
7 </level>
8
9 <level type="theme" level="2">
10  .
11  <table name="theme-a" rads="..">
12    <item left_append="a" features="ipfv" var="a" />
13  </table>
14  .
15  .
16 </level>
17
18 <level type="exponent" level="3">
19  .
20  <item block="1" prefix="n" features="1.ipfv" />
21  .
22  .
23  .
24 </level>

```

FIGURE 5.14 : La définition des niveaux réalisationnels dans Alexina<sub>PARSL</sub> (données de MaltLex)

### 5.2.2.3 Représentation de la supplétion dans Alexina<sub>PARSL</sub>

Le formalisme Alexina d'origine n'avait pas de moyen pour représenter les cas de supplétion de forme. Les deux manières dont la supplétion de forme pouvait être simulée revenait aux deux premières solutions évoquées plus haut pour simuler l'allomorphie de radical :

1. Lister les formes allomorphiques dans le fichier contenant les formes ajoutées manuellement et ne pas les générer avec la grammaire. Cette solution requiert d'associer par ailleurs les entrées lexicales concernées par ce type d'allomorphie à des classes flexionnelles défectives pour les cases comportant de l'allomorphie.
2. Associer les entrées supplétives à des classes flexionnelles dans lesquelles le radical de l'entrée en question est considéré comme un segment vide (ou quasi-vide) et les formes complètes sont modélisées comme des suffixes s'accolant à ce radical vide. La forme entière est alors traitée comme la désinence correspondant à une case donnée. Cette solution est notamment employée pour la flexion du verbe ÊTRE dans le *Lefff*.

Dans Alexina<sub>PARSLI</sub>, il est possible de lister les formes supplétives qui se substituent à la formation régulière par la grammaire dans le lexique. Pour la forme de la première personne du pluriel de l'indicatif présent du verbe ÊTRE en français, nous pouvons ainsi directement spécifier la forme *sommes* dans le lexique. Cette forme se substitue à une forme régulière qui comporterait la désinence régulière *ons* du 1PL.PRS.IND.

Les formes supplétives sont ainsi indiquées à la fin des spécifications morphologiques lexicales dans le fichier comportant les entrées intensionnelles. Nous donnons un exemple de supplétion de forme pour le verbe DIRE du français à la figure 5.15. Ce verbe a la particularité de former son 2PL.PRS.IND par la forme supplétive *dites*.

```
1 dire v:3re/dis,,di,,,,,,,,,dit/2.pl.prs.ind=dites
```

FIGURE 5.15 : La supplétion de formes dans Alexina<sub>PARSLI</sub> (données du *Lefff* modifié)

En revanche, dans le cas de formes supplétives surabondantes, il suffit de lister les formes en question dans le fichier comportant les formes supplémentaires listées manuellement. Nous donnons un exemple de cela pour la flexion de l'adjectif MARRON en français. MARRON est listé dans le *Lefff* comme étant un adjectif ne variant qu'en nombre et étant invariable en genre, cf. la figure 5.16. Par ailleurs, le *Lefff* comporte une forme de pluriel invariable supplémentaire dans le fichier des formes manuelles, cf. la figure 5.17. La forme (*marron y* est indiquée avec son lemme (*marron*) et l'étiquette morphosyntaxique exprimée (*p* pour pluriel)).

```
1 marron adj-2
```

FIGURE 5.16 : La supplétion surabondante de formes dans Alexina<sub>PARSLI</sub> : entrée intensionnelle standard (données du *Lefff* d'origine)

```
1 marron marron p
```

FIGURE 5.17 : La supplétion surabondante de formes dans Alexina<sub>PARSLI</sub> : forme supplétive supplémentaire (données du *Lefff* original)

#### 5.2.2.4 La définition des zones dans Alexina<sub>PARSLI</sub>

Une des nouveautés principales dans Alexina<sub>PARSLI</sub> est l'introduction de la notion de zones telle que définie dans *PARSLI*. Dans Alexina<sub>PARSLI</sub>, elle apparaît dans les schèmes.

Les zones peuvent être définies de deux façons différentes : soit directement par des balises indiquant une zone à un niveau donné, cf. la figure 5.18, soit par la définition des espaces partitionnants à l'intérieur d'un niveau, cf. la figure 5.23 page 302, soit en combinant l'indication d'un espace partitionnant dans un schème et son intersection avec une classe flexionnelle, cf. la figure 5.19. Les exemples que nous donnons aux figures 5.18 à 5.23 sont deux façons de représenter des zones du niveau exposant de la flexion verbale latine selon notre analyse proposée au chapitre 1.

```

1 <zone name="A3" rads="..">
2
3 <item block="1-2" append="re" features="prs.inf" />
4
5 <item block="1-2" append="isse" features="pst.inf" />
6
7 <item block="1-2" append="ns" features="nom.sg.prs.ptcp" />
8
9 <item block="1" append="nt" features="prs.ptcp" />
10 <item block="1" append="ur" features="fut.ptcp" />
11 <item block="1" append="" features="sup" />
12 <item block="1" append="nd" features="grv" />
13
14 <like block="2" liketablename="adj-oa" features="fut.ptcp|sup" />
15 <like block="2" liketablename="adj-oa" likefeatures="acc.sg.n|gen.
    sg.n|abl.sg.n" features="grv" />
16 <like block="2" liketablename="adj-c" features="prs.ptcp|grv" />
17
18 </zone>

```

FIGURE 5.18 : La définition (directe) des zones dans Alexina<sup>PARSLI</sup> (données de *Leffla* modifié)

```

1 <pattern name="aAB-reg" >
2   <subpattern>
3     <zone level="1" table="s-reg"/>
4     <zone level="2" table="a"/>
5     <zone level="3" table="v-B"/>
6     <zone level="3" partitionspace="I3" table="v-A"/>
7   </subpattern>
8 </pattern>

```

FIGURE 5.19 : La définition des zones dans les schèmes d'Alexina<sup>PARSLI</sup> (données de *Leffla*)

### 5.2.2.5 La définition des classes flexionnelles dans Alexina<sub>PARSLI</sub>

En Alexina<sub>PARSLI</sub>, les classes réalisationnelles peuvent être représentées directement par les balises `table` comme dans le formalisme Alexina d'origine. Dans ce cas les règles de réalisation seront directement indiquées à l'intérieur des balises `table` en question<sup>26</sup>. Elle peuvent néanmoins également être indiquées par des référence à des zones indépendantes, comme illustré par l'exemple figure 5.20 page 300. Dans ce cas, des balises faisant référence à ces zones sont situées à l'intérieur de balises `table` définissant la classe réalisationnelle en question.

### 5.2.2.6 La définition des espaces partitionnants dans Alexina<sub>PARSLI</sub>

Dans <sub>PARSLI</sub> nous avons introduit la notion d'espace partitionnant comme des sous-ensembles d'ensembles de paquets de traits morphosyntaxiques pouvant être exprimés par une classe d'entrée lexicales donnée. Ces espaces partitionnant sont un des moyens de faire référence aux zones réalisationnelles telles que définies dans <sub>PARSLI</sub>. Dans Alexina<sub>PARSLI</sub>, ces espaces partitionnants peuvent être spécifiés pour chaque niveau. Nous illustrons le mécanisme à la figure 5.21 page 301 pour le niveau radical (données de MaltLex) et à la figure 5.22 page 301 pour le niveau exposant (données de *Leffa*).

La figure 5.23 page 302 illustre la définition d'une zone réalisationnelle par l'intermédiaire d'un espace partitionnant. Cette zone est employée par le schème flexionnel à la figure 5.19 page 298.

### 5.2.2.7 La représentation des schèmes dans Alexina<sub>PARSLI</sub>

Dans Alexina, les entrées intensionnelles sont associées à des classes flexionnelles. Alexina<sub>PARSLI</sub> reprend la conception de <sub>PARSLI</sub> selon laquelle une entrée lexicale est associée à un ensemble de zones à travers un schème. Les classes flexionnelles ne sont qu'une notion dérivée des zones : elles émergent des généralisations observables sur les zones communément associées conjointement à un ensemble significatif d'entrées lexicales. Les classes flexionnelles sont indiquées par les balises `table` dans Alexina<sub>PARSLI</sub>, les zones par les balises `zone` ou, indirectement, par la définition des espaces partitionnants `partitionspace`.

Ainsi, les entrées lexicales du lexique intensionnel sont-elles associées à des schèmes, cf. les entrées du latin extraites de *Leffa* et indiquées à la figure 5.24 page 303. Les schèmes associés aux entrées sont par ailleurs définis dans la grammaire à travers les zones employées

26. Ou à défaut à l'intérieur de balises `zone`, elles mêmes situées à l'intérieur des balises `table`.

```

1  <!-- Table definitions at level 3 -->
2
3  <table name="v-ae" >
4    <zoneref name="v-ae1" />
5    <zoneref name="v-ae-ic2" />
6  </table>
7
8  <table name="v-ic" >
9    <zoneref name="v-ic1" />
10   <zoneref name="v-ae-ic2" />
11 </table>
12
13 <!-- Zone definitions at level 3 -->
14
15 <zone name="v-ae1">
16   <item block="1" append="ba" features="pst.ipfv.ind" />
17   <item block="1" append="bi" features="fut.ipfv.ind" />
18
19   <item block="1" append="ns" features="nom.sg.prs.ptcp|acc.sg.prs.
20     ptcp.n" />
21   <item block="1" append="nt" features="prs.ptcp" />
22   <item block="1" append="nd" features="grv|grd" />
23 </zone>
24
25 <zone name="v-ae-ic2">
26   <item block="1" append="" features="prs.ipfv.ind|imp|prs.inf" />
27   <item block="1" append="a" features="prs.ipfv.sbjv" />
28   <item block="1" append="re" features="pst.ipfv.sbjv" />
29
30   <item block="1" append="i" features="prs.pfv.ind" except="3.pl"/>
31   <item block="1" append="era" features="pst.pfv.ind" />
32   <item block="1" append="isse" features="pst.pfv.sbjv" />
33   <item block="1" append="eri" features="fut.pfv.ind|prs.pfv.sbjv" />
34
35   <item block="1" append="ur" features="fut.ptcp" />
36   <item block="1" append="" features="inf|sup|pst.ptcp|3.pl.prs.pfv.
37     ind" />
38 </zone>

```

FIGURE 5.20 : La définition de classes réalisationnelles par référence à des zones indépendantes dans Alexina<sub>PARSL</sub> (données de *Leffa*)

```

1 <level type="stem" level="1">
2   <partitionspace name="S1" features="3.m.sg.pfv"/>
3   <partitionspace name="S2" features="1.pfv|2.pfv"/>
4   <partitionspace name="S3" features="3.f.sg.pfv"/>
5   <partitionspace name="S4" features="3.pl.pfv"/>
6   <partitionspace name="S5" features="sg.ipfv"/>
7   <partitionspace name="S6" features="pl.ipfv"/>
8
9   <partitionspace name="3PFV" features="3.pfv"/>
10  <partitionspace name="CH" except="3.pfv|pl.ipfv"/>

```

FIGURE 5.21 : La définition des espaces partitionnants dans Alexina<sub>PARSLI</sub> au niveau radical (données de MaltLex)

```

1 <level type="exponent" level="3">
2   <partitionspace name="I1" features="ipfv.ind|ipfv.sbjv|prs.inf"/>
3   <partitionspace name="I2" features="pfv.ind|pfv.sbjv|pst.inf"/>
4   <partitionspace name="I3" features="prs.ptcp|fut.ptcp|fut.inf|sup|
   pst.ptcp|grv|grd"/>

```

FIGURE 5.22 : La définition des espaces partitionnants dans Alexina<sub>PARSLI</sub> au niveau exposant (données de Leffa)

à chaque niveau, comme illustré à la figure 5.25 page 304. Toutefois<sup>27</sup> lorsqu'un schème ne spécifie ni zone ni espace partitionnant mais uniquement une classe flexionnelle (`table=`), cette dernière est interprétée par le programme de flexion comme une zone unique ayant pour propriété secondaire de constituer une classe flexionnelle à elle seule, comme indiqué à la figure 5.25.

Dans la grammaire flexionnelle, les schèmes associés aux entrées lexicales intensionnelles spécifient, outre les zones et le nom des schèmes, la catégorie flexionnelle de l'entrée associée au schème en question (`cat=`) et une règle de transfert (optionnelle lorsqu'il s'agit de l'identité). Spécifier la catégorie flexionnelle détermine notamment la liste des cases devant être réalisées par le paradigme. Cette information est ainsi en interaction directe avec la nouvelle façon de définir les étiquettes morphosyntaxiques et les cases des paradigmes flexionnels dans Alexina<sub>PARSLI</sub>. À la figure 5.25, nous illustrons la représentation des schèmes dans la grammaire à l'aide d'exemples de schèmes extraits de notre implémentation de la flexion verbale du latin décrite au chapitre 1.

La spécification de `on_failure="skip"` permet de simplifier l'implémentation en spécifiant deux zones d'un même niveau dans un même sous-schéme `subpattern` pour

27. Notamment pour des raisons de rétrocompatibilité avec Alexina où les entrées étaient directement associées à des classes flexionnelles.



```

1 <level type="exponent" level="3">
2   <partitionspace name="I1" features="ipfv.ind|ipfv.sbjv|prs.inf"/>
3   <partitionspace name="I2" features="pfv.ind|pfv.sbjv|pst.inf"/>
4   <partitionspace name="I3" features="prs.ptcp|fut.ptcp|fut.inf|sup|pst.
      ptcp|grv|grd"/>
5
6 <table name="v-A" rads=".(a|e)" exclude="ptcp.pst|grd">
7   <item block="1-2" append="e" features="2.sg.prs.ipfv.imp" var="e-imp"
      />
8   <item block="1-2" append="i" features="2.sg.prs.ipfv.imp" var="i-imp"
      />
9   <item block="1-2" append="" features="2.sg.prs.ipfv.imp" />
10  <item block="1-2" append="re" features="prs.inf" />
11  <item block="1-2" append="nto" features="3.pl.fut.ipfv.imp" />
12  <item block="1-2" append="isse" features="pst.inf" />
13  <item block="1-2" append="i" features="1.sg.prs.pfv.ind" />
14  <item block="1-2" append="erunt" features="3.pl.prs.pfv.ind" />
15  <item block="1-2" append="isti" features="2.sg.prs.pfv.ind" />
16  <item block="1-2" append="istis" features="2.pl.prs.pfv.ind" />
17  <item block="1-2" append="ns" features="nom.sg.prs.ptcp" />
18
19  <item block="1" append="a" features="prs.ipfv.sbjv" />
20  <item block="1" append="ba" features="pst.ipfv.ind" />
21  <item block="1" append="re" features="pst.ipfv.sbjv" />
22  <item block="1" append="bi" features="fut.ipfv.ind" />
23  <item block="1" append="" features="prs.ipfv.ind|prs.ipfv.imp" />
24  <item block="1" append="to" features="fut.ipfv.imp" />
25  <item block="1" append="i" features="prs.pfv.ind" />
26  <item block="1" append="era" features="pst.pfv.ind" />
27  <item block="1" append="isse" features="pst.pfv.sbjv" />
28  <item block="1" append="eri" features="fut.pfv.ind|prs.pfv.sbjv" />
29  <item block="1" append="nt" features="prs.ptcp" />
30  <item block="1" append="ur" features="fut.ptcp" />
31  <item block="1" append="" features="sup" />
32  <item block="1" append="nd" features="grv" />
33
34  <item block="2" append="o" features="1.sg.ind" except="1.sg.pst.ind"/>
35  <item block="2" append="m" features="1.sg.sbjv|1.sg.pst.ind" />
36  <item block="2" append="s" features="2.sg" except="imp" />
37  <item block="2" append="t" features="3.sg" except="imp" />
38  <item block="2" append="mus" features="1.pl" />
39  <item block="2" append="tis" features="2.pl" except="imp" />
40  <item block="2" append="nt" features="3.pl" except="imp" />
41  <item block="2" append="te" features="2.pl.ipfv.imp" />
42  <like block="2" liketablename="adj-oa" features="fut.ptcp|sup" />
43  <like block="2" liketablename="adj-oa" likefeatures="acc.sg.n|gen.sg.n
      |abl.sg.n" features="grv" />
44  <like block="2" liketablename="adj-c" features="prs.ptcp|grv" />
45 </table>
46 ...
47 </level>

```

FIGURE 5.23 : La définition des zones par les espaces partitionnants dans Alexina<sup>PARSL</sup> (données de Leffla)

```

1 celo v-aA
2 celor v-aB
3 conor v-aAB

```

FIGURE 5.24 : La représentation des schèmes dans les entrées intensionnelles d'Alexina<sub>PARSLI</sub> (données de Leffa)

peu que les zones en questions soient associées à des traits strictement distincts. Dans l'exemple de la figure 5.25, la spécification de `on_failure="skip"` est indiquée pour les `realzone` du niveau réalisationnel 4. Cette spécification garantit que des formes sont produites quand bien même aucune des zones de niveau 4 ne s'appliquent, à savoir pour l'ensemble formes finies et les infinitifs.

### 5.2.2.8 Les règles de transfert dans Alexina<sub>PARSLI</sub>

Les schèmes définis dans la grammaire peuvent également comporter la mention des règles de transfert définissant les zones réalisationnelles associées, pour chaque niveau défini, à chaque entrée lexicale flexionnelle dans  $\mathcal{PARSLI}$ . La mention de ces règles de transfert est optionnelle dans l'implémentation en Alexina<sub>PARSLI</sub> d'une formalisation en  $\mathcal{PARSLI}$  dès lors que la règle de transfert en question est la fonction identité, cf. les schèmes de la flexion latine indiqués à la figure 5.25. En revanche, lorsque la fonction de transfert est différente de la fonction identité, elle doit être mentionnée pour les zones réalisationnelles exprimant un décalage morphosyntaxique dans les schèmes des verbes concernés. Les données du serbo-croate que nous avons décrites au chapitre 3 puis réévoquées au chapitre 4 peuvent ainsi être implémentées par leur association au schème de la figure 5.27 page 306. Rappelons que les noms du croate dont nous parlons, et que nous rappelons au tableau 5.6, ont la propriété de former leur pluriel en empruntant la zone du singulier d'une classe flexionnelle distincte (cf. tableau 5.5 page 305) de celle à laquelle ils sont par ailleurs associés. Ils sont ainsi à la fois hétéroclites et décalés. La fonction de transfert qui les définit au niveau exposant a ainsi la propriété d'associer à une fonction de pluriel exprimée l'expression morphologique du singulier :  $t(PL) = sg$ . Cette fonction de transfert est indiquée dans la définition de la zone correspondante au niveau exposant dans une implémentation en Alexina<sub>PARSLI</sub>, cf. les figures 5.26 page 305 et 5.27 page 306. Ce mécanisme nécessite par ailleurs une définition de la fonction de transfert en question dans la section concernant les traits morphosyntaxiques dans la grammaire flexionnelle : en l'occurrence nous définissons une fonction de transfert `transfer` appelée `sg2pl` et qui a pour caractéristique de transformer la valeur `sg` du trait `NUMBER` en la valeur `pl`.

```

1 <pattern name="v-aA" cat="v">
2   <subpattern>
3     <realzone level="1" table="s-reg"/>
4     <realzone level="2" table="a"/>
5     <realzone level="3" table="v-A" />
6     <realzone level="4" table="adj-oa" features="ptcp.fut|grv"
7       on_failure="skip" />
8     <realzone level="4" table="adj-c" features="ptcp.prs" on_failure="
9       skip"/>
10   </subpattern>
11 </pattern>
12
13 <pattern name="v-aB" cat="v">
14   <subpattern>
15     <realzone level="1" table="s-reg"/>
16     <realzone level="2" table="a"/>
17     <realzone level="3" table="v-B" />
18     <realzone level="4" table="adj-oa" features="ptcp.pst|grd"
19       on_failure="skip" />
20   </subpattern>
21 </pattern>
22
23 <pattern name="v-aAB" cat="v">
24   <subpattern>
25     <realzone level="1" table="s-reg"/>
26     <realzone level="2" table="a"/>
27     <realzone level="3" table="v-B" />
28     <realzone level="4" table="adj-oa" features="ptcp.pst|grd"
29       on_failure="skip" />
30   </subpattern>
31   <subpattern>
32     <realzone level="1" table="s-reg"/>
33     <realzone level="2" table="a"/>
34     <realzone level="3" partitionspace="I3" table="v-A" />
35     <realzone level="4" table="adj-oa" features="ptcp.fut|grv"
36       on_failure="skip" />
37     <realzone level="4" table="adj-c" features="ptcp.prs" on_failure="
38       skip"/>
39   </subpattern>
40 </pattern>

```

FIGURE 5.25 : La représentation des schèmes dans les entrées intensionnelles d'Alexina<sub>PARSL</sub> (données de Leffa)

	$Z_{N-et}^{exp}$ : NEUTRE EN -ET DUGME « bouton »		$Z_{M-a}^{exp}$ : MASCULIN JELEN « cerf »		$Z_{F-a}^{exp}$ : FÉMININ EN -A ŽENA « femme »	
	$z_{N-et,sg}^{exp}$ : SG	$z_{N-et,pl}^{exp}$ : PL	$z_{M,sg}^{exp}$ : SG	$z_{M,pl}^{exp}$ : PL	$z_{F-a,sg}^{exp}$ : SG	$z_{F-a,pl}^{exp}$ : PL
NOM	<i>dugme</i>	<i>dugm-et-a</i>	<i>jelen</i>	<i>jelen-i</i>	<i>žen-a</i>	<i>žen-e</i>
ACC	<i>dugme</i>	<i>dugm-et-a</i>	<i>jelen-a</i>	<i>jelen-e</i>	<i>žen-u</i>	<i>žen-e</i>
GEN	<i>dugm-et-a</i>	<i>dugm-et-a</i>	<i>jelen-a</i>	<i>jelen-a</i>	<i>žen-e</i>	<i>žen-a</i>
DAT	<i>dugm-et-u</i>	<i>dugm-et-ima</i>	<i>jelen-u</i>	<i>jelen-ima</i>	<i>žen-i</i>	<i>žen-ama</i>
LOC	<i>dugm-et-u</i>	<i>dugm-et-ima</i>	<i>jelen-u</i>	<i>jelen-ima</i>	<i>žen-i</i>	<i>žen-ama</i>
INS	<i>dugm-et-om</i>	<i>dugm-et-ima</i>	<i>jelen-om</i>	<i>jelen-ima</i>	<i>žen-om</i>	<i>žen-ama</i>

TABLE 5.5 : Déclinaison régulière de noms en serbo-croate

	NEUTRE EN -ET/ RADICAL EN -A DETE « enfant »		MASCULIN / RADICAL EN -A BRAT « frère »	
	$z_{N-et,sg}^{exp}$ : SG	$z_{F-a,sg}^{exp}$ : PL	$z_{M,sg}^{exp}$ : SG	$z_{F-a,sg}^{exp}$ : PL
NOM	<i>dete</i>	<i>dec-a</i>	<i>brat</i>	<i>brać-a</i>
ACC	<i>dete</i>	<i>dec-u</i>	<i>brat-a</i>	<i>brać-u</i>
GEN	<i>det-et-a</i>	<i>dec-e</i>	<i>brat-a</i>	<i>brać-e</i>
DAT	<i>det-et-u</i>	<i>dec-i</i>	<i>brat-u</i>	<i>brać-i</i>
LOC	<i>det-et-u</i>	<i>dec-i</i>	<i>brat-u</i>	<i>brać-i</i>
INS	<i>det-et-om</i>	<i>dec-om</i>	<i>brat-om</i>	<i>brać-om</i>

TABLE 5.6 : Décalage dans deux paradigmes de noms en serbo-croate

Dans le lexique :

- 1 *dugme* n-Net
- 2 *jelen* n-M
- 3 *žena* n-Fa
- 4 *dete* n-Net-Fa
- 5 *brat* n-M-Fa

FIGURE 5.26 : Entrées lexicales intensionnelles pour les noms du serbo-croate en Alexina<sub>PARSLI</sub>

Dans la grammaire :

```

1  <!-- MORPHOSYNTACTIC FEATURES -->
2
3  <!-- features -->
4
5  <mfsbundle name="NUMBER" features="sg|pl"/>
6
7  <!-- transfer rules -->
8      <!-- transfer=id is default -->
9
10 <transfer name="sg2pl" source="pl" target="sg"/>
11
12
13 <!-- patterns -->
14
15 <pattern name="n-Net" cat="n" >
16     <subpattern>
17         <realzone level="1" table="s-reg"/>
18         <realzone level="2" table="n-Net"/>
19     </subpattern>
20 </pattern>
21
22 <pattern name="n-Ma" cat="n" >
23     <subpattern>
24         <realzone level="1" table="s-reg"/>
25         <realzone level="2" table="n-Ma"/>
26     </subpattern>
27 </pattern>
28
29 <pattern name="n-Fa" cat="n" >
30     <subpattern>
31         <realzone level="1" table="s-reg"/>
32         <realzone level="2" table="n-Fa"/>
33     </subpattern>
34 </pattern>
35
36 <pattern name="n-Net-Fa" cat="n" >
37     <subpattern>
38         <realzone level="1" table="s-reg"/>
39         <realzone level="2" zone="n-Net-sg" />
40         <realzone level="2" zone="n-Fa-sg" transfer="sg2pl"/>
41     </subpattern>
42 </pattern>
43
44 <pattern name="n-Ma-Fa" cat="n" >
45     <subpattern>
46         <realzone level="1" table="s-reg"/>
47         <realzone level="2" zone="n-Ma-sg" />
48         <realzone level="2" zone="n-Fa-sg" transfer="sg2pl"/>
49     </subpattern>
50 </pattern>

```

FIGURE 5.27 : La représentation des fonctions de transfert dans les schèmes d'Alexina PARS<sub>L</sub>

5.2.2.9 Les nouvelles opérations morphologiques dans Alexina<sup>PARSLI</sup>

Dans le formalisme Alexina d'origine, les seules opérations disponibles pour exprimer des règles de réalisation étaient des opérations de préfixation ou de suffixation. Toute opération morphologique plus complexe nécessitait dans l'ancienne version d'Alexina d'utiliser les règles morphographémiques/morphonologiques de façon détournée. Alexina<sup>PARSLI</sup> permet désormais d'exprimer des opérations morphologiques non concaténatives, notamment par la définition dans la grammaire des opérations employées par la morphologie d'une langue donnée. Ces dernières sont en effet ainsi considérées comme faisant partie de la particularité du système morphologique d'une langue donnée. La suffixation (notée par `append`) et la préfixation (notée par `left_append`) restent cependant définies par défaut dans Alexina<sup>PARSLI</sup>. À ces deux opérations par défaut s'ajoutent par ailleurs les opérations `insert` permettant l'insertion d'un segment et `replace` permettant la substitution d'un segment par un autre. Ces opérations de base peuvent être utilisées pour définir des opérations plus complexes telles celles des figures 5.29 à 5.31<sup>28</sup>.

La définition d'opérations repose le plus souvent sur la définition de classes de graphèmes (ou digraphes) ou de phonèmes, cf. la définition des classes de graphèmes en maltais extraite de MaltLex à la figure 5.28.

```

1 <letterclass name="C" letters="b c d f g g ħ ħ h j k l m n p q r s t
   v w x z z ' "/>
2 <letterclass name="B" letters="għ ħ h q ' "/>
3 <letterclass name="V" letters="a e i o u ie"/>
4 <letterclass name="R" letters="għ l m n r"/>
5 <letterclass name="nR" letters="b c d f g g ħ h j k p q s t v w x z z
   ' "/>

```

FIGURE 5.28 : Classes de graphèmes/phonèmes Alexina<sup>PARSLI</sup> (données de MaltLex)

À partir des classes de segments définies et des opérations primitives (`append`, `left_append`, `insert` et `replace`), des opérations comme celles de la figure 5.29 page 308 peuvent ainsi être définies. L'opération `deleteV1` définie pour les radicaux du maltais indique par exemple que pour une source (`source`) de construction de radical ayant une structure CVCVC l'application de `deleteV1` génère un nouveau radical (`target`) à la structure CCVC. Une opération `replace` copie tout segment appartenant à la classe de segment C telle que définie à la figure 5.28 en ce même segment s'il est marqué par `[ : C : ]` autant dans la source que dans la sortie. Si un segment est noté avec un 0 dans

28. Notons que ces opérations sont formellement une redéfinition d'expressions régulières par une syntaxe plus proche de celle habituellement utilisée dans les descriptions linguistiques.

la source, comme c'est le cas de  $[0:V:]$  dans l'opération `deleteV1`, alors le segment en question n'est pas copié dans la sortie et s'en trouve effacé.

```

1 <operation_definition name="deleteV1">
2 <replace source = "[[:C:][0:V:][[:C:][[:V:][[:C:]]" target="[:C:][[:C:][[:V:
   :][[:C:]]" rev="-" />
3 <replace source = "[[:C:][0:V:][[:C:][[:V:]]" target="[:C:][[:C:][[:V:]]" rev=
   "-" />
4 </operation_definition>

```

FIGURE 5.29 : Opérations morphologiques en Alexina<sub>PARSLI</sub> (données de MaltLex)

Par ailleurs, il n'est pas seulement possible d'effacer des segments avec l'opération primitive `replace`. Ne pas copier un segment marqué par 0 dans la source peut également être l'occasion de le remplacer par un autre segment, marqué par `_` dans la sortie. L'opération en question sera alors une opération comportant des arguments<sup>29</sup>, cf. la figure 5.30 page 308. Dans les opérations utilisées dans les zones de réalisation, l'opération `deleteV1changeV2` définie à la figure 5.30 sera ainsi indiquée `deleteV1changeV2 (i)` si la deuxième voyelle est remplacée par un « *i* ».

```

1 <operation_definition name="deleteV1changeV2">
2 <replace source = "[[:C:][0:V:][[:C:][0:V:][[:C:]]" target="[:C:][[:C:]]_[:C
   :]" rev="-" />
3 <replace source = "[[:C:][0:V:][[:C:][0:V:]]" target="[:C:][[:C:]]_" rev="-"
   />
4 </operation_definition>

```

FIGURE 5.30 : Opérations morphologiques avec argument(s) en Alexina<sub>PARSLI</sub> (MaltLex)

Enfin, la copie de segments, telle que nécessaire pour la reduplication, est également possible dans le nouveau formalisme Alexina<sub>PARSLI</sub>. La figure 5.31 page 309 montre ainsi la définition d'une opération de reduplication pour la construction de certains radicaux verbaux du latin. La consonne initiale *y* est recopiée deux fois, l'identité des deux consonnes réécrites étant représentée par l'indice 1 dans  $[1:C:]$  dans le formalisme. On note par ailleurs que l'opération définie à la figure 5.31 page 309 est une opération à deux variables : la voyelle d'origine ( $[0:V]$ ) est effacée et remplacée par un segment `_` dont l'identité, ainsi que celle de l'autre segment supplémentaire marqué par le second `_`, sera donnée par les arguments indiqués dans la règle de réalisation `redup-initial (x, y)`.

29. Autant que de `_` dans la définition de l'opération.

```

1 <operation_definition name="redup-initial">
2 <replace source = "#[:C:][0:V:]" target="#[1:C:]_[1:C:]_"/>
3 </operation_definition>

```

FIGURE 5.31 : Opérations morphologiques avec argument(s) en Alexina<sub>PARSLI</sub> (Leffla)

### 5.2.2.10 Les traits morphosyntaxiques et la définition des cases des paradigmes dans Alexina<sub>PARSLI</sub>

Dans le formalisme Alexina d'origine, les traits morphosyntaxiques n'apparaissent que sous la formes d'étiquettes associées aux règles de réalisation, cf. l'exemple des règles de réalisation extraites du *Lefff* original à la figure 5.11 page 290. Alexina<sub>PARSLI</sub> introduit une gestion des traits morphosyntaxiques en cohérence avec leur statut au sein du modèle <sub>PARSLI</sub>.

<sub>PARSLI</sub> est un modèle réalisationnel. Chaque forme *y* est conçue comme la réalisation d'un paquets de traits morphosyntaxiques. Son implémentation en Alexina<sub>PARSLI</sub> emploie la notion de STRUCTURE DE TRAITS.

Le système de calcul des cases à réaliser pour une catégorie donnée repose sur le système de l'unification des structures de traits en Alexina<sub>PARSLI</sub>. Pour une catégorie flexionnelle donnée (*category*), les cases à réaliser sont calculées comme l'unification des traits exprimables non incompatibles. Dans l'exemple des traits du maltais donnés à la figure 5.32 page 310, les cases à réaliser sont calculées par le croisement des traits *mode*, *perfectivity*, *number*, *person* et *gender* indiqués comme existant pour la catégorie flexionnelle *v* (verbe). Chaque case comporte donc, à ce stade, une valeur pour l'un de ces traits. Les valeurs possibles sont indiquées par *features=* pour chaque trait (*mfsbundle*). Ainsi *pfv* et *ipfv* sont les deux valeurs possibles du trait *perfectivity*.

Toutefois, certaines valeurs, voire certains traits dans leur ensemble, peuvent être incompatibles avec certaines autres valeurs ou traits. Dans le cas du maltais illustré par la figure 5.32, le mode impératif *imp* ne peut ainsi être réalisé en même temps que l'aspect perfectif *pfv*. De même, il ne sera pas associé à une valeur du trait *gender*, l'impératif ne faisant pas de distinction de genre en maltais. Afin de calculer le bon inventaire de cases réalisées par une catégorie flexionnelle donnée, Alexina<sub>PARSLI</sub> possède ainsi un mécanisme supplémentaire d'exclusion de la concomitance de certains traits ou de certaines valeurs. Dans notre exemple de la figure 5.32, ce mécanisme, représenté par la balise *mfsexclusion*, permet ainsi d'exclure la case \*3SG.F.PFV de l'inventaire des cases possibles — quand bien même cette dernière aurait pu être réalisée au vu des seules



définitions de traits et de leurs valeurs données par les balises `mfsbundle`.

```

1 <!-- MORPHOSYNTACTIC FEATURES -->
2
3 <category name="v">
4   <mfsbundle name="mode" features="unmarked|imp"/>
5   <mfsbundle name="perfectivity" features="pfv|ipfv"/>
6   <mfsbundle name="number" features="sg|pl"/>
7   <mfsbundle name="person" features="1|2|3"/>
8   <mfsbundle name="gender" features="m|f"/>
9
10  <mfsexclusion features="pfv" excludes="standard"/>
11  <mfsexclusion features="imp" excludes="pfv"/>
12  <mfsexclusion features="imp" excludes="gender"/>
13  <mfsexclusion features="imp" excludes="person"/>
14  <mfsexclusion features="1|2" excludes="gender"/>
15  <mfsexclusion features="pl" excludes="gender"/>
16 </category>

```

FIGURE 5.32 : La définition des cases à réaliser dans Alexina $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$  (données de MaltLex)

Une autre façon de spécifier des cases interdites dans un système flexionnel donné est de lister les cases en question<sup>30</sup> par des balises `invalidfeatureset`, cf. l'exemple du latin à la figure 5.33 page 310 excluant les cases comportant la spécification `*INF.IPFV.PST` de l'inventaire des cases possibles.

```

1 <invalidfeatureset features="inf.ipfv.pst"/>

```

FIGURE 5.33 : Définition de la défektivité (données de *Leffa*)

Enfin, Alexina $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$  permet également de représenter par des spécifications sur les traits les types de défektivité définis pour un système flexionnel donné par les balises `featureset`. Dans l'exemple du latin que nous donnons à la figure 5.34 page 311, les verbes défectifs à l'imperfectif marqués dans le lexique par un type de défektivité décrit dans la grammaire par `featureset name="ipfv"` sont indiqués comme ayant pour seules cases réalisables les cases compatibles (c'est-à-dire unifiables) avec les structures de traits spécifiant les valeurs `ipfv.ind`, `ipfv.sbjv`, `ptcp`, `grd` et `grv`. De même, les verbes défectifs à l'infinitif futur et pour l'adjectif verbal, le gérondif et le supin seront spécifiés par un inventaire de paquets de traits morphosyntaxiques défektivité défini dans la grammaire par `<featureset name="no-fut-iffin"` comme réalisant toutes les cases possibles

30. Des cases qui devraient sinon être produites en raison des traits définis par les balises `<mfsbundle`.

(`exist=""`) selon les balises `mfsbundle` sauf celles unifiales avec les structures de traits spécifiant les valeurs `fut.imp`, `grd`, `grv`, `sup`.

```

1 <featureset name="ipfv" exist="ipfv.ind|ipfv.sbjv|ptcp|grd|grv"/>
2 <featureset name="impers" exist="3.ind|sbjv.|.ptcp|grd|grv"/>
3 <featureset name="impers-sg" exist="3.sg.ind|3.sg.sbjv|ptcp|grd|grv|
  inf"/>
4 <featureset name="impers-sg-ipfv" exist="3.sg.ipfv.ind|3.sg.ipfv.sbjv|
  ptcp|grd|grv"/>
5
6 <featureset name="no-fut-iffin" exist="" except="fut.imp|grd|grv|sup"/>
7 <featureset name="no-fut-ptcp" exist="" except="fut.ptcp|sup"/>

```

FIGURE 5.34 : Définition de la défektivité (données de *Leffla*)

### 5.2.2.11 La représentation des règles de réalisation dans *Alexina*<sub>PARSLI</sub>

Dans le formalisme *Alexina* d'origine, les opérations possibles pour chaque règle de réalisation étaient limitées à la préfixation et la suffixation. L'introduction d'un mécanisme de définition d'opérations morphologiques propres à une langue dans *Alexina*<sub>PARSLI</sub> a permis de rendre compte de cas plus complexes et notamment de traiter les cas de morphologie non concaténative.

Par ailleurs, dans le formalisme *Alexina* d'origine, chaque règle de réalisation exprimait à chaque fois une étiquette morphosyntaxique complète par l'affixation sur un radical *a priori* invariable lui même. Dans *Alexina*<sub>PARSLI</sub>, l'introduction des niveaux permet d'exprimer des règles de réalisation successives réalisant en un premier temps tous les traits liés à la sélection d'un radical, puis avec la sélection d'une zone thématique et enfin des traits effectivement exprimés par les exposants d'une forme donnée. Or, chaque réalisation opérée à un niveau supplémentaire doit être compatible avec les traits réalisés dans les niveaux précédents. Cette compatibilité est garantie par un calcul sur l'unification des traits réalisés par chacune des règles successives. Si l'unification réussit, la règle est appliquée, si elle échoue, elle ne l'est pas.

Plus précisément, à l'intérieur d'un niveau donné, une règle de réalisation ou règle d'exponence, en *Alexina*<sub>PARSLI</sub> repose pareillement sur le mécanisme d'unification. Nous pouvons ainsi, comme en PFM, définir des blocs de règles à l'intérieur des niveaux qui réalisent successivement un certain nombre de paquets de traits morphosyntaxiques. Dans chaque bloc, une et une seule règle s'applique pour remplir une case donnée ; au bloc suivant, une règle ne s'applique qu'à condition que le paquet de traits morphosyntaxiques qu'elle exprime corresponde à ce qui est attendu par cette case, c'est-à-dire si elle exprime

une structure de traits unifiable avec celles exprimées par les règles appliquées aux blocs précédents.

Dans l'exemple que nous donnons pour les exposants du maltais à la figure 5.35 page 312, la réalisation du niveau exposant est prise en compte par deux blocs de règles `block="1"` et `block="2"`, le premier réalisant l'aspect et la personne, le second le marquage ou non du pluriel (`suffix="u" features="pl"`). Comme en PFM, Alexina<sub>PARSLI</sub> permet par ailleurs d'écrire des règles porte-manteau ayant précedence sur les règles par blocs standard. L'étendue des porte-manteaux y est indiquée par les blocs couverts, dans notre exemple `block="1-2"` permet de court-circuiter les deux blocs `block="1"` et `block="2"` définis plus haut.

Notons enfin qu'à ce stade, s'il y a compétition entre deux règles d'un même niveau (ou à défaut d'un même bloc) pour la réalisation d'un ensemble de traits, la première règle indiquée sera réalisée. Alexina<sub>PARSLI</sub> diffère du fonctionnement de PFM où il doit être possible de calculer la règle la plus spécifique<sup>31</sup> afin de sélectionner la règle à appliquer. Dans son état actuel, Alexina<sub>PARSLI</sub> implémente donc un ordre intrinsèque des règles au sein d'un bloc, à la manière de EWP (Anderson, 1992), plutôt que de s'appuyer sur le principe de Pañini (Stump, 2001).

```

1 <level type="exponent" level="3">
2
3   <table name="exponence" rads="">
4
5     <item block="1-2" suffix="na" features="1.pl.pfv"/>
6     <item block="1-2" suffix="et" features="3.f.sg.pfv"/>
7
8     <item block="1" suffix="t" features="1.sg.pfv|2.pfv"/>
9
10    <item block="1" prefix="n" features="1.ipfv"/>
11    <item block="1" prefix="t" features="2.ipfv"/>
12    <item block="1" prefix="t" features="3.f.sg.ipfv"/>
13    <item block="1" prefix="j" features="3.ipfv"/>
14
15    <item block="2" suffix="u" features="pl"/>
16
17  </table>
18
19 </level>

```

FIGURE 5.35 : La représentation des règles de réalisation dans Alexina<sub>PARSLI</sub> (données de MaltLex)

31. En termes de traits réalisés dans un bloc donné.

## 5.3 Implémentation d'une description du système verbal du maltais

L'objet de cette section est de donner un exemple d'implémentation d'une analyse morphologique  $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$  et également de montrer la pertinence de l'implémentation pour la garantie de la cohérence et de l'exhaustivité de l'analyse. L'analyse morphologique que nous donnons comme exemple porte sur une partie du système verbal maltais. Elle repose sur l'analyse de l'allomorphie radicale de ce sous-ensemble du système verbal à base sémitique proposé par Camilleri (2011). Cette analyse rend compte de la majeure partie des alternances de radicaux observables pour ce sous-ensemble de verbes, mais elle laisse de côté la réalisation complète des formes. Dans l'analyse que nous avons proposée dans (Camilleri & Walther, 2012), nous présentons une modélisation des formes complètes dans  $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$  ainsi qu'une implémentation de cette analyse en  $\text{Alexina}_{\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}}$ . Le lexique associé à cette analyse comporte des données extraites du *Maltese Language Resource Server* (Spagnol, 2011). L'implémentation sur un échantillon quasi-exhaustif<sup>32</sup> du sous-ensemble de verbes traités nous a permis de compléter l'analyse de Camilleri (2011) par les cas qui n'étaient pas couverts par son analyse. Elle met également l'accent sur une complexité particulière dans la construction des formes complètes liées à la sélection de la voyelle d'extension portée par certains verbes à l'imperfectif. Cette caractéristique dépassait la sélection de radicaux au sens strict telle qu'analysée par Camilleri (2011). Elle n'était ainsi pas non plus couverte par son analyse. Sa prise en compte découle directement des exigences liées à la formalisation complète des formes d'une part et de l'implémentation pour la quasi-exhaustivité des verbes concernés de l'autre.

### 5.3.1 La flexion verbale maltaise

#### 5.3.1.1 Le maltais

Le MALTAIS est une langue parlée par environ 300 000 locuteurs. Comrie et al. (2009) indiquent que, génétiquement, le maltais appartient à la branche sud-arabique du sémitique central. Il se caractérise notamment par la combinaison de sa grammaire sémitique avec d'importantes influences romanes, et plus particulièrement italiennes. En particulier, le lexique maltais est plus empreint d'influences romanes que sémitiques. En termes dialectologiques, le maltais comporte ainsi un substrat arabe sicilien/maghrébin, un superstrat roman et, suite à la conquête et présence anglaise sur l'île entre 1800 et 1964,

32. 560 lemmes pour les seuls verbes appartenant au premier binyan des verbes maltais à base sémitique.

un adstrat anglais.

Notre étude porte sur une sous-partie du système verbal du *maltais standard* (MS). Le système morphologique des langues sémitiques est traditionnellement divisé en plusieurs sous-classes selon les racines consonnantiques sous-jacentes. Ces sous-classes, en général verbales ou nominales, appelées BINYAN dans la tradition linguistique sémitique, sont reliées les unes aux autres par des relations dérivationnelles, telle que la dérivation d'un verbe causatif à partir d'un autre verbe<sup>33</sup>. Notre étude traite ici le premier binyan du MS, c'est-à-dire le binyan comportant les lexèmes verbaux considérés comme non dérivés. Notons que la sélection de ce binyan pour une première analyse du sous-système verbal maltais à base sémitique n'est pas arbitraire puisqu'il correspond à plus de 50% des lexèmes verbaux à base sémitique du MS, ainsi qu'aux mots forme les plus fréquents en corpus<sup>34</sup>.

### 5.3.1.2 Paradigmes verbaux et classes flexionnelles

La flexion verbale en MS est relativement simple si l'on s'en tient à la seule variation des marques flexionnelles affixales. Nous adoptons donc la stratégie employée par Boyé (2000) et les travaux qui en découlent et qui consiste à postuler une seule classe flexionnelle pour rendre compte de l'intégralité des paradigmes verbaux<sup>35</sup>. Cette stratégie maximisant la flexion constante entre les différents types de verbes est employée par Boyé pour rendre compte de la flexion verbale dans les langues romanes. Nous illustrons la classe flexionnelle unique pour les verbes du MS en bleu au tableau 5.7 page 315. La forme de citation employée traditionnellement en maltais est celle de la troisième personne du masculin singulier du perfectif, cf. KITEB « écrire », BENA « construire », SAR<sup>36</sup> « devenir », LAQAT « frapper » et HATAF « saisir » dans le tableau 5.8 page 316.

En revanche, la complexité de la construction des formes s'accroît dès lors que l'on observe la variation des radicaux au sein du premier binyan. Les radicaux au sein d'un même paradigme varient par rapport à la forme de citation sur deux points : d'une part en ce qu'ils conservent ou non l'une ou les deux de leurs voyelles radicales, et en ce que la

---

33. Pour une étude de la dérivation verbale en maltais, voire notamment Hoberman & Aronoff (2002).

34. Par ailleurs, Ussishkin & Twist (2009) montrent à l'aide d'expériences psycholinguistiques que le premier binyan comporte également les formes les plus régulières au sens où des tests de choix lexicaux visuels autant qu'auditifs ont produit des temps de réaction plus bas pour les lexèmes du premier binyan que pour ceux des autres.

35. Une stratégie alternative consiste à postuler l'existence d'un nombre plus important de classes flexionnelles, maximisant ainsi la régularité des radicaux verbe par verbe. Pour le maltais, cette seconde stratégie, que nous ne suivons pas ici, est notamment employée par Mifsud (1995) et Fabri (2009).

36. Dans les exemples de formes, nous ajoutons dans notre notation le symbole API « : » pour marquer la longueur des voyelles, quand bien même la notation suit par ailleurs les conventions graphémiques du maltais qui ne comportent pas de marquage de la longueur vocalique.

PFV	1SG		RADICAL	-t
	2SG		RADICAL	-t
	3SG.M		RADICAL	
	3SG.F		RADICAL	-et
	1PL		RADICAL	-na
	2PL		RADICAL	-tu
	3PL		RADICAL	-u
IPFV	1SG	n-(v)	RADICAL	
	2SG	t-(v)	RADICAL	
	3SG.M	j-(v)	RADICAL	
	3SG.F	t-(v)	RADICAL	
	1PL	n-(v)	RADICAL	-u
	2PL	t-(v)	RADICAL	-u
	3PL	j-(v)	RADICAL	-u

TABLE 5.7 : Marques affixales portées par les verbes du premier binyan

voyelle en question est modifiée ou non ; d'autre part en un changement systématique dans le placement de l'accent. Comme le montre le tableau 5.8, les radicaux verbaux peuvent ainsi afficher jusqu'à six formes différentes ; ce cas extrême est ici illustré par le verbe BENA « construire ». Par ailleurs, la répartition de syncrétismes entre les radicaux d'un même paradigme produit, d'un paradigme à l'autre, des schèmes de radicaux différents, fondement de l'analyse proposée par Camilleri (2011, 2012) et que nous décrivons à la section 5.3.2.

Enfin, certains verbes affichent également une voyelle d'extension entre le préfixe personnel et le radical à l'imperfectif. Cette voyelle est marquée en vert dans le tableau 5.7 et en gras pour les verbes concernés dans le tableau 5.8.

Pour les lexèmes les plus allomorphiques, les paradigmes affichent ainsi un total de six radicaux, répartis comme indiqué dans le tableau 5.9.

### 5.3.1.3 Remarques sur les analyses précédentes

Une très grande partie des analyses de la flexion maltaise existantes (cf. ci-dessous) s'appuient sur la tradition sémitique définissant le comportement morphologique des lexèmes à partir de la structure de leur racine consonantique sous-jacente, en l'occurrence au tableau 5.8 les racines marquées entre parenthèses pour chaque verbe, comme la racine K-T-B pour KITEB « écrire ». De telles approches sont notamment proposées par McCarthy (1981) et Lowenstamm (1996) pour l'arabe standard et Mifsud (1995) et

	KITEB « écrire » (K-T-B)	BENA « construire » (B-N-J)	SAR « devenir » (S-J-R)	LAQAT « frapper » (L-Q-T)	HATAF « saisir » (H-T-F)	MESS « toucher » (M-S-S)
PFV 1SG	<i>ktib-t</i>	<i>bnej-t</i>	<i>sir-t</i>	<i>lqat-t</i>	<i>htaf-t</i>	<i>messej-t</i>
PFV 2SG	<i>ktib-t</i>	<i>bnej-t</i>	<i>sir-t</i>	<i>lqat-t</i>	<i>htaf-t</i>	<i>messej-t</i>
PFV 3SG.M	<i>kiteb</i>	<i>bena</i>	<i>sa:r</i>	<i>laqat</i>	<i>hataf</i>	<i>mess</i>
PFV 3SG.F	<i>kitb-et</i>	<i>bnie-t</i>	<i>sa:r-et</i>	<i>laqt-et</i>	<i>hatf-et</i>	<i>mess-et</i>
PFV 1PL	<i>ktib-na</i>	<i>bnej-na</i>	<i>sir-na</i>	<i>lqat-na</i>	<i>htaf-na</i>	<i>messej-na</i>
PFV 2PL	<i>ktib-t-u</i>	<i>bnej-t-u</i>	<i>sir-t-u</i>	<i>lqat-t-u</i>	<i>htaf-t-u</i>	<i>messej-tu</i>
PFV 3PL	<i>kitb-u</i>	<i>bn-e:w</i>	<i>sa:r-u</i>	<i>laqt-u</i>	<i>hatf-u</i>	<i>messe:-w</i>
IPFV 1SG	<i>n-i-kteb</i>	<i>n-i-bni</i>	<i>n-si:r</i>	<i>n-o-lqot</i>	<i>n-a-htaf</i>	<i>n-miss</i>
IPFV 2SG	<i>t-i-kteb</i>	<i>t-i-bni</i>	<i>s-si:r</i>	<i>t-o-lqot</i>	<i>t-a-htaf</i>	<i>t-miss</i>
IPFV 3SG.M	<i>j-i-kteb</i>	<i>j-i-bni</i>	<i>j-si:r</i>	<i>j-o-lqot</i>	<i>j-a-htaf</i>	<i>j-miss</i>
IPFV 3SG.F	<i>t-i-kteb</i>	<i>t-i-bni</i>	<i>s-si:r</i>	<i>t-o-lqot</i>	<i>t-a-htaf</i>	<i>t-miss</i>
IPFV 1PL	<i>n-i-ktb-u</i>	<i>n-i-bn-u</i>	<i>n-si:r-u</i>	<i>n-o-lqt-u</i>	<i>n-a-htf-u</i>	<i>n-miss-u</i>
IPFV 2PL	<i>t-i-ktb-u</i>	<i>t-i-bn-u</i>	<i>s-si:r-u</i>	<i>t-o-lqt-u</i>	<i>t-a-htf-u</i>	<i>t-miss-u</i>
IPFV 3PL	<i>j-i-ktb-u</i>	<i>j-i-bn-u</i>	<i>j-si:r-u</i>	<i>j-o-lqt-u</i>	<i>j-a-htf-u</i>	<i>j-miss-u</i>
IMP 2SG	<i>i-kteb</i>	<i>i-bni</i>	<i>si:r</i>	<i>o-lqot</i>	<i>a-htaf</i>	<i>miss</i>
IMP 2PL	<i>i-ktb-u</i>	<i>i-bn-u</i>	<i>si:r-u</i>	<i>o-lqt-u</i>	<i>a-htf-u</i>	<i>miss-u</i>

TABLE 5.8 : Paradigmes de verbes du premier binyan

	PERFECTIVE	IMPERFECTIVE
1SG	RAD2 -t	n- RAD5
2SG	RAD2 -t	t- RAD5
3SGM	RAD1	j- RAD5
3SGF	RAD3 -et	t- RAD5
1PL	RAD1 -na	n- RAD6 -u
2PL	RAD1 -tu	t- RAD6 -u
3PL	RAD4 -u	j- RAD6 -u

TABLE 5.9 : Différentiation maximale et répartition des radicaux dans un paradigme maltais

Fabri (2009) pour le maltais. Ces analyses considèrent que chaque racine consonantique comporte trois ou quatre consonnes, appelées RADICAUX<sup>37</sup>. Elles établissent ainsi une distinction entre verbes FORTS, SOURDS et FAIBLES reposant sur l'idée d'une structure sous-jacente de la structure consonantique. Ces trois types de verbes sont représentés

37. Ce n'est pas cet emploi du terme *radical* que nous utilisons dans les autres sections de cette thèse.

respectivement par KITEB « *écrire* », MESS « *toucher* » et SAR « *devenir* ». En particulier, la classe des verbes faibles se caractérise par la présence de consonnes dites faibles, à savoir les consonnes *w* et *j*<sup>38</sup>. On compte trois types de verbes faibles, en fonction de la place de la consonne faible dans structure sous-jacente : les verbes à consonne faible initiale (ou assimilative) comme dans WIRET « *hériter* » (**W**-R-T), médiane (ou creuse) comme pour SAR « *aller* » (M-**J**-R) ou encore finale (ou manquante/défective) comme dans BENA « *construire* » (B-N-**J**). Dans cette tradition, on appelle habituellement BASE VERBALE les différents types verbaux existants. Ainsi, on distingue :

1. le type fort de structure sous-jacente C-C-C, où C est n'importe quelle consonne forte,
2. le type faible initial **w**-C-C, où **w** représente une consonne faible,
3. le type géminé C-C<sub>i</sub>-C<sub>i</sub> comportant deux consonnes identiques et ainsi de suite.

Camilleri montre que la postulation d'une racine consonantique sous-jacente n'est pas nécessaire pour décrire le comportement des verbes maltais, voire ne permet justement pas de rendre compte de façon satisfaisante des alternances de radicaux observées dans les paradigmes des verbes du premier binyan. Parmi les arguments de Camilleri figurent notamment les points suivants :

1. Les analyses à racines consonantiques reflètent des distinctions diachroniquement valides, mais n'ont pas de réalité dans le système linguistique accessible aux locuteurs.
2. Elles nécessitent de stipuler l'existence de verbes irréguliers par leur racine consonantique sous-jacente alors même que ces verbes peuvent être traités de façon parfaitement régulière dans une approche s'appuyant sur des radicaux concrets.
3. Les régularités et irrégularités exprimées par des analyses à racines consonantiques ne correspondent pas nécessairement aux régularités et irrégularités observables en surface, alors que les analyses reposant sur la notion de radical reflètent (nécessairement) plus fidèlement la réalité observable.
4. Connaître la structure consonantique sous-jacente ne suffit pas à construire les paradigmes verbaux dans la mesure où certains verbes irréguliers affichent le même type de structure consonantique mais construisent leurs formes différemment dans leurs paradigmes de surface : ainsi les verbes LAQAT « *frapper* » et HATAF « *saisir* » partagent une structure consonantique forte (C-C-C) mais n'affichent pas le même nombre de radicaux différents dans leurs paradigmes respectifs. Les alternances de radicaux sont représentées par les changements de couleur au sein de chaque paradigme dans le tableau 5.8.

38. Les autres consonnes sont dites « *fortes* ».



À ces quatre arguments nous en ajoutons trois supplémentaires portant sur les formes fléchies complètes et non seulement sur les radicaux étudiés par Camilleri.

5. Dans le sous-paradigme de l'imperfectif, certains verbes affichent une voyelle d'extension (ou thématique) entre le radical et les préfixes personnels de l'imperfectif (*n-*, *t-* et *j-*). Cette extension n'apparaît que pour les verbes dont le radical de surface à la troisième personne du singulier masculin est de structure CVCVC, cf. nos exemples KITEB « écrire », LAQAT « frapper » et HATAF « saisir », ou CVCV comme pour BENA « construire » au tableau 5.8. La présence de cette voyelle d'extension est systématique pour ces deux types de verbes, quelle que soit leur structure consonantique sous-jacente (consonnes fortes ou faibles). Elle n'apparaît pas pour les autres verbes. Ainsi, dans le tableau 5.8, SAR (S-J-R) « devenir » et MESS (M-S-S) « toucher » quoique de structures consonantiques différentes<sup>39</sup>, ont en commun le fait de ne présenter aucune voyelle thématique entre le préfixe de l'imperfectif et le radical. Ce fait peut en revanche se déduire de leur appartenance aux verbes à structures CV:C et CVCC respectivement. Cette caractéristique est également partagée par les verbes à structure CVCCVC.

Traiter tous les verbes, indépendamment de leur structure de surface à la forme de citation (3SG.F.PFV) et uniformément comme des verbes tri- ou quadri-consonantiques sous-jacents, ne permet pas de distinguer les verbes à structure de surface CVCVC et CVCV des verbes ayant une autre structure de surface comme CV:C, ni de ce fait de prédire l'apparition de la voyelle d'extension à l'imperfectif pour les seuls verbes à structure CVCVC et CVCV. Ce n'est que depuis la structure de surface de la forme de citation que la présence de la voyelle d'extension peut être prédite.

6. La voyelle d'extension n'est par ailleurs pas la même pour tous les verbes : soit *i*, soit *e*, soit *a*, soit *o*, selon les verbes concernés.

Dans nos exemples du tableau 5.8 nous trouvons ainsi un *-i-* à l'imperfectif des verbes KITEB « écrire » et BENA « construire », un *-o-* pour LAQAT « frapper » et un *-a-* pour HATAF « saisir ».

Une première approximation serait de dire que la voyelle d'extension est *i* dans le cas général, mais s'ajuste par harmonie vocalique à la voyelle du radical de l'imperfectif lorsque la consonne initiale est une consonne arrière. Ainsi on prédit bien la voyelle d'extension *i* pour BENA et KITEB comme le cas général, ainsi que *a* pour le lexème à consonne arrière initiale HATAF. En revanche, dans le cas de LAQAT, on s'attendrait

39. Dans le classement traditionnel, SAR est un verbe trilitaire à médiane faible tandis que MESS est un verbe géminé.

à des formes à voyelle d'extension *i* de type *\*nilqot* pour la première personne du singulier de l'imperfectif, le *l* n'étant pas une consonne arrière. Pourtant, ici aussi l'harmonie vocalique opère et la forme obtenue est *nolqot*.

Ni la structure consonantique sous-jacente, ni même la nature de la première consonne, ne permettent de prédire la couleur de la voyelle d'extension des verbes en CVCVC et CVCV

7. Enfin, dans la mesure où une forme de surface (celle du 3SG.M.PFV) permet de rendre compte des propriétés paradigmatiques des verbes maltais, il semble inutilement compliqué de vouloir recourir à une forme sous-jacente non attestée pour les décrire.

### 5.3.2 Analyse des verbes maltais selon Camilleri (2011)

Camilleri (2011, 2012) propose une analyse du système verbal maltais qui repose sur les seules alternances de radicaux au sein des paradigmes. Elle fait l'observation que certaines cases des paradigmes maltais sont systématiquement syncrétiques entre elles quant aux radicaux qu'elles comportent : à savoir les première et deuxième personnes (singulier et pluriel) du perfectif, et le singulier *vs.* le pluriel à l'imperfectif. Il est ainsi possible de regrouper les cases systématiquement syncrétiques en six groupements de radicaux, résumés au tableau 5.10. Ces six groupements de radicaux peuvent varier ou ne pas varier entre eux, modifiant ainsi le nombre de syncrétismes entre les radicaux d'un paradigme donné. Nous les indiquons de 1 à 6.

ENSEMBLE	CASES CONCERNÉES	
RAD2	PFV	1SG, 2SG, 1PL, 2PL
RAD1	PFV	3SG.M
RAD3	PFV	3SG.F
RAD4	PFV	3PL
RAD5	IPFV	SG
RAD6	IPFV	PL

TABLE 5.10 : Cases à radicaux systématiquement syncrétiques

Camilleri regroupe par ailleurs les différents verbes selon la structure de leur forme de citation, à savoir la troisième personne du singulier masculin du perfectif. Ils peuvent ainsi appartenir à l'une des cinq grandes classes suivantes : CVCVC, CVCCVC, CVCV, CVCC ou CV:C. C'est pour ces macro-classes qu'elle emploie le terme de BASE VERBALE (*verbal base*).

Chacune de ces bases verbales comporte un certain nombre de lexèmes (la base verbale CVCVC étant de loin la plus peuplée). Chaque lexème affiche à son tour des alternances de radicaux dont la structure consonne/voyelle est régulière d'un paradigme à l'autre au sein d'une même base verbale. Ainsi, les verbes appartenant à la base verbale CVCVC présentent tous une structure CVCVC à la troisième personne du singulier masculin du perfectif, c'est-à-dire dans pour le RAD1 défini dans le tableau 5.10. Ils affichent cependant des structures CCVC dans pour le RAD2 et le RAD5 et CVCC dans pour le RAD3 et le RAD4. Pour le RAD6 leur structure dépend de la nature de la deuxième consonne. La structure associée au RAD6 est CCC si la consonne centrale est une consonne sonante (à savoir l'une des consonnes *gh*, *l*, *m*, *n* ou *r*) et CVCC dans les autres cas. Nous représentons cette ambiguïté par une structure C(V)CC dans le tableau 5.11. Ce tableau indique également les variations de structure affichées par les autres bases verbales, ainsi que pour chaque forme la voyelle accentuée (marquée *́*) et les frontières de syllabes.

Deux radicaux d'un même verbe ayant la même structure peuvent se distinguer par des alternances vocaliques. Ainsi, le verbe KITEB « *écrire* » dans le tableau 5.8 possède en plus de son radical CVCVC de la forme de citation (RAD1), deux radicaux à structure CCVC (*ktib* et *kteb*) pour le RAD2 et le RAD5 respectivement) et un radical à structure CVCC (*kitb* pour le RAD2 et le RAD4) et un radical à structure CCC (*ktb* pour le RAD6). KITEB illustre le fait que pour une même structure CCVC (*ktib* et *kteb*), un changement vocalique (*e/i*) entraîne l'existence de deux radicaux différents pour les deux groupements de radicaux concernés (RAD2 et RAD4), tandis que l'absence de changement vocalique entre les deux radicaux à structure CVCC (*kitb*) provoque la présence d'un syncrétisme entre les radicaux le RAD2 et le RAD4.

Compte tenu de ces observations, Camilleri classe ensuite, pour chaque base verbale, les verbes en fonction de la distribution des syncrétismes de radicaux qu'ils affichent entre les six groupements de radicaux disponibles dans leurs paradigmes. Elle obtient des sous-classes syncrétiques pour chaque base verbale, indiquées par les couleurs du tableau 5.8. Les syncrétismes supplémentaires qui affectent ainsi systématiquement certains des groupements de radicaux définis dans le tableau 5.10 permettent d'établir, de façon orthogonale aux bases verbales, des classes de distribution syncrétique des radicaux, appelées *stem pattern classes* par Camilleri (2011). Nous utiliserons en français le terme de CLASSES DE SYNCRÉTISME RADICAL (CSR). Le verbe BENA « *construire* » qui affiche six radicaux différents dans son paradigme correspondra ainsi à la CSR ABCD EF, chaque lettre représentant un radical différent correspondant respectivement aux groupements de radicaux RAD1 à RAD6 dans le tableau 5.10. Comme indiqué au tableau 5.12, le RAD1

	CVVCVC	CVCCVC	CVCV	CVCC	CV:C	
PFV	1SG	CCV́Ct	CCV́jt	CVC.Céjt	CV́Ct	
	2SG	CCV́Ct	CCV́jt	CVC.Céjt	CV́Ct	
	1PL	CCV́C.na	CCV́j.na	CVC.Céj.na	CV́C.na	
	2PL	CCV́C.tu	CCV́j.tu	CVCCéj.tu	CV́C.tu	
	3.M.SG	CV́.CVC	CV́.CV	CV́CC	CV́:C	
	3.F.SG	CV́C.Cet	CCV́:t/CCít	CV́C.Cet	CV́:.Cet	
	3PL	CV́C.Cu	CCV́w	CVCCéw	CV́:.Cu	
	IPFV	1SG	nV́C.CVC	nV́C.CV	nCV́CC	nCV́:C
		2SG	tV́C.CVC	tV́C.CV	tcV́CC	tcV́:C
3.M.SG		jV́C.CVC	jV́C.CV	jcV́CC	jcV́:c	
3.F.SG		tV́C.CVC	nV́C.CV	tcV́CC	tcV́:C	
1PL		nV́C.CCu/nV́.CV́C.Cu	nV́C.Cu/nVC.CV́w	nCV́C.Cu	nCV́:.Cu	
2PL		tV́C.CCu/tV́.CV́C.Cu	tV́C.Cu/tVC.CV́w	tcV́C.Cu	tcV́:.Cu	
3PL		jV́C.CCu/jV́.CV́C.Cu	jV́C.Cu/jVC.CV́w	jcV́C.Cu	jcV́:.Cu	
IMP		SG	V́C.CVC	V́C.CV	CV́CC	CV́:C
		PL	V́C.CCu/V́.CV́C.Cu	v.CCv́w/V́C.Cu	CV́CC	CV́:C

TABLE 5.11 : Structures des radicaux dans les paradigmes du premier binyan

correspond ainsi au radical A chez Camilleri (2011).

ENSEMBLE	CASES CONCERNÉES		
RAD2	A	PFV	1SG, 2SG, 1PL, 2PL
RAD1	B	PFV	3SG.M
RAD3	C	PFV	3SG.F
RAD4	D	PFV	3PL
RAD5	E	IPFV	SG
RAD6	F	IPFV	PL

TABLE 5.12 : Notation de la différenciation maximale dans les classes de syncrétisme radical (CSR) dans (Camilleri, 2011)

Notons par ailleurs que des classes verbales différentes peuvent participer de la même CSR, ce qui est le cas pour la CSR ABCC DC qui appartient à la fois aux bases verbales CVCVC et CVCV. Nous illustrons la répartition des CSR par rapport aux classes verbales à la figure 5.36. Le graphique indique ainsi pour le premier binyan (nœud racine en haut du graphique, niveau 0), les classes verbales existantes (niveau 1 sous le nœud racine), et pour chaque classe verbale, la CSR qu'elle peut implémenter (niveau 2). Les CSR peuvent par ailleurs être regroupées en superclasses de syncrétisme (niveau 3) selon les syncrétismes observables au perfectif.

Les changements vocaliques sont définis pour chaque groupement de radicaux et pour chaque base verbale en fonction des CSR et des voyelles de la forme de citation. Camilleri (2011) indique un certain nombre de changements vocaliques possibles pour chaque base verbale. Nous listons ces changements vocaliques dans les tableaux 5.13 à 5.15<sup>40</sup>.

Chacun de ces tableaux indique, pour une classe verbale donnée, les CSR existantes (en colonne)<sup>41</sup>. Sont également indiqués, ligne par ligne, les schèmes vocaliques existants : par exemple, pour la classe verbale CVCVC, les voyelles peuvent être telles qu'elles donnent lieu aux schèmes suivants : CaCaC, CaCeC, CeCeC, CeCaC et CiCeC et CoCoC. À l'image de ce qui est proposé par Camilleri (2011), nous noterons ces schèmes vocaliques par respectivement A-A, A-E, E-E, E-A, I-E et O-O. Ils définissent les lignes dans les tableaux 5.13 à 5.15.

Certains schèmes auront par ailleurs pour deuxième consonne une consonne sonnante (notée R) ou une consonne non sonnante (notée  $\bar{R}$ ). Comme le montrent les tableaux 5.13

40. L'implémentation liée à notre analyse des formes complètes décrite à la section 5.3.3 nous a amenée à compléter, voire à corriger, ce premier inventaire des changements vocaliques indiqué par Camilleri (2011).

41. Nous n'avons pas indiqué les exemples isolés de verbes complètement irréguliers comme HA « prendre » ou ĠIE « venir » qui méritent d'être traités séparément dans la description du système flexionnel du maltais.



à 5.15, cette caractéristique aura également une influence sur les changements vocaliques affectant les verbes à base verbale CVCVC.

Chacun de ces schèmes peut, au cours de la formation des radicaux, subir des changements vocaliques. Ils sont alors indiqués sous le schème correspondant pour le groupement radical concerné. Ainsi, pour le schème CaCaC appartenant à la CSR ABCC AE, le premier *a* de la forme de citation devient *e* lors de la formation du RAD6. La notation employée pour exprimer ce changement dans le schéma est la suivante.

5.1      RAD6 : V<sub>1</sub> a → e

L'intégralité des possibilités indiquées par Camilleri (2011) est ainsi résumée dans les cases dans différentes nuances de rouge des tableaux 5.13 à 5.15. Les cases non remplies correspondent à des cas de figure non attestés.

Notons que Camilleri n'indique pas l'intégralité des changements vocaliques effectivement observables dans les paradigmes du premier binyan. Parmi les verbes à base verbale CVCC, elle indique en particulier les changements pour les verbes comme HASS « *ressentir* » dont la voyelle de la forme de citation est « *a* », mais pas ceux des verbes comme MESS « *toucher* » dont la voyelle de la forme de citation est « *e* ». Notre travail d'implémentation nous a amenée à ajouter les alternances valables pour les verbes de ce type ainsi que pour les autres verbes manquants dans l'inventaire des types d'alternance de (Camilleri, 2011). Nous indiquons ces ajouts en teintes de bleu dans les tableaux 5.13 à 5.15.

Pour les verbes à base verbale CVCC, Camilleri indique que le radical est étendu à la fin par l'ajout de *ej* ou de *e:* respectivement pour les radicaux RAD2 et RAD4. Cette analyse se démarque des analyses plus standard (Mifsud, 1995 ; Fabri, 2009) qui posent ici l'existence d'une classe flexionnelle à part contenant les séquences *ej* et *e:*. Camilleri propose ainsi ici une segmentation différente.

Dans le tableau 5.14, nous notons l'extension finale du radical par les deux règles indiquées ci-dessous :

5.2      RAD1 : +ej  
           RAD4 : +e:

Elles provoquent ainsi pour un radical comme celui de HASS « *ressentir* » la construction des radicaux RAD2 *ħassej* et RAD4 *ħasse:*.

Les verbes à base verbale CVCCVC sont quant à eux entièrement absents de l'étude dans

CVCVC		ABCC DE	ABCC AE	ABCC DC	ABCC AC
A-A	$\bar{R}$	RAD5 $V_2 a \rightarrow o$	— —		
	R	RAD5 $V_2 a \rightarrow o$ RAD6 $V_1 a \rightarrow o$	RAD6 $V_1 a \rightarrow e$		— —
	R	RAD5 $V_2 a \rightarrow i$ RAD6 $V_1 a \rightarrow o$			
A-E	$\bar{R}$	RAD2 $V_2 e \rightarrow i$			
	R	RAD6 $V_1 a \rightarrow i$			
E-E	$\bar{R}$	RAD2 $V_2 e \rightarrow i$	RAD2 $V_2 e \rightarrow i$ RAD5 $V_2 e \rightarrow i$		
	R	RAD2 $V_2 e \rightarrow i$ RAD6 $V_1 e \rightarrow i$			
E-A	$\bar{R}$	RAD5 $V_2 a \rightarrow o$	— —		
	R	RAD5 $V_2 a \rightarrow o$ RAD6 $V_1 e \rightarrow o$	RAD6 $V_1 e \rightarrow i$		
	R				
I-E	$\bar{R}$	RAD2 $V_2 e \rightarrow i$			
	R			RAD2 $V_2 e \rightarrow i$	
O-O	$\bar{R}$		— —		
	R				— —

TABLE 5.13 : Changements vocaliques pour les verbes à base CVCVC

CVCC	ABBC DD	ABBB DD
A	RAD2 +ej	RAD2 +ej
	RAD4 +e:	RAD4 +e:
	RAD5 $V_1 a \rightarrow o$	RAD5 $V_1 a \rightarrow o$
	RAD6 $V_1 a \rightarrow o$	RAD6 $V_1 a \rightarrow o$
E	RAD2 +ej	RAD2 +ej
	RAD4 +e:	RAD4 +e:
	RAD5 $V_1 e \rightarrow i$	RAD5 $V_1 e \rightarrow i$
	RAD6 $V_1 e \rightarrow i$	RAD6 $V_1 e \rightarrow i$
I	RAD2 +ej	
	RAD4 +e:	

TABLE 5.14 : Changements vocaliques pour les verbes à base CVCC



CVC	ABBB CC (type 1)	ABBB CC (type 2)	ABBB CD
A	RAD2 V <sub>2</sub> a: → o RAD6 V <sub>1</sub> a: → u:	RAD2 V <sub>2</sub> a: → i RAD6 V <sub>1</sub> a: → i:	
IE	RAD2 V <sub>2</sub> ie → o RAD6 V <sub>1</sub> ie → u:	RAD2 V <sub>2</sub> ie → i RAD6 V <sub>1</sub> ie → i:	RAD2 V <sub>2</sub> ie → o RAD6 V <sub>1</sub> ie → u:

TABLE 5.15 : Changements vocaliques pour les verbes à base CVC

CVCV	ABCD EF	ABCD ED	ABCD DC
E-A	RAD2 +ej RAD3 V <sub>2</sub> a → ie RAD4 V <sub>2</sub> a → e: RAD5 V <sub>2</sub> a → i	RAD2 +ej RAD3 V <sub>2</sub> a → ie RAD4 V <sub>2</sub> a → e: RAD6 V <sub>2</sub> a → e:	
A-A	RAD2 +ej RAD3 V <sub>2</sub> a → ie RAD4 V <sub>2</sub> a → e: RAD5 V <sub>2</sub> a → i	RAD2 +ej RAD3 V <sub>2</sub> a → ie RAD4 V <sub>2</sub> a → e: RAD4 V <sub>2</sub> a → e:	RAD2 +aj RAD3 V <sub>2</sub> a → a:
O-A	RAD2 +ej RAD3 V <sub>2</sub> a → ie RAD4 V <sub>2</sub> a → e: RAD5 V <sub>2</sub> a → i		
I-A		RAD2 +ej RAD3 V <sub>2</sub> a → ie RAD4 V <sub>2</sub> a → e: RAD6 V <sub>2</sub> a → e:	

TABLE 5.16 : Changements vocaliques pour les verbes à base CVCV

(Camilleri, 2011). Il nous a ainsi fallu établir l'intégralité de leurs propriétés pour notre analyse décrite à la section 5.3.3.

CVCCVC	ABCC BC	BBCC BC
A-A		— —
A-E	RAD2 V <sub>2</sub> e → i	
E-E	RAD2 V <sub>2</sub> e → i	
E-A		— —
I-E	RAD2 V <sub>2</sub> e → i	
O-O		— —

TABLE 5.17 : Changements vocaliques pour les verbes à base CVCCVC

Connaître la CSR et la forme de citation (qui renseigne directement sur la base verbale) d'un lexème suffit ainsi presque toujours pour construire l'intégralité des radicaux de son paradigme. Une exception importante est toutefois le cas des lexèmes à base verbale CV:C dont la forme de citation est de type Ca:C ou CieC. Les changements vocaliques les affectant peuvent être de deux types pour les Ca:C et de trois types pour les CieC<sup>42</sup>, cf. le tableau 5.18. Une autre exception sont les radicaux en CaCaC des CSR ABCC DE qui peuvent se distinguer dans leur comportement à l'imperfectif, tel qu'indiqué au tableau 5.13.

ABBB CC	Ca:C	Ca:C	CieC	CieC	CieC
RAD2	ciC	CoC	ciC	CoC	ciC
RAD1	Ca:C	Ca:C	CieC	CieC	CieC
RAD3 =RAD1	Ca:C	Ca:C	CieC	CieC	CieC
RAD4 =RAD1	Ca:C	Ca:C	CieC	CieC	CieC
RAD5	ci:C	cu:C	ci:C	cu:C	cu:C
RAD6 =RAD5	ci:C	cu:C	ci:C	cu:C	cu:C

TABLE 5.18 : Changements vocaliques dans les CSR du premier binyan

Chaque lexème est ainsi associé à une CSR particulière. Par exemple, BENA « construire », est un verbe CVCV appartenant à la CSR ABCD EF. Il a donc un radical différent dans chacun des groupements de radicaux disponibles et subit les changements vocaliques suivant :

RAD2 : structure CCV où *a* devient *aj* : *bnaj*.

RAD1 : structure CVCV sans changement : *bena*.

RAD3 : structure CCV où *a* devient *ie* : *bnie*.

RAD4 : structure CCV où *a* devient *e* : *bnē*.

RAD5 : structure CCV sans changement : *bna*.

RAD6 : structure CC sans changement : *bn*.

L'appartenance d'un lexème donné à une CSR particulière est décrite par Camilleri (2011) comme étant une caractéristique purement morphologique au sens de Corbett & Baerman (2006), qui détermine ainsi la structure du paradigme du lexème en question. Elle décrit ainsi les CSR comme les résultats de fonctions morphomiques (Aronoff, 1994)

42. Autant pour les verbes en Ca:C qu'en CieC, on note toutefois une corrélation entre la forme RAD2 en ciC et la forme RAD5 en ci:C d'une part et la forme RAD2 en CoC et la forme RAD5 en cu:C de l'autre. Le troisième type des verbes en CieC peut ainsi être considéré comme une combinaison des deux premiers, c'est-à-dire comme un cas d'hétéroclise entre deux CSR.

affectant de façon systématique, mais non motivée en dehors de la morphologie, des classes particulières de lexèmes.

### 5.3.3 Analyse des formes complètes dans $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}\mathcal{I}$

Nous avons formalisé l'analyse de Camilleri (2011) présentée ci-dessus et implémenté cette formalisation en Alexina $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}\mathcal{I}$ . Le lexique associé à notre implémentation repose sur les 850 verbes du premier binyan du *Maltese Language Resource Server* de Spagnol (2011). Comme ces derniers comprenaient un bon nombre de lemmes incorrects, ils ont été triés au préalable pour ne garder que les 600 verbes (environ) appartenant effectivement au premier binyan en maltais. Notre formalisation de l'analyse de Camilleri (2011) nous a permis d'évaluer la cohérence de sa proposition, mais également d'en montrer les limites, sur lesquelles nous revenons ci-dessous. Elle a également étendu l'analyse qui jusqu'ici ne portait que sur les syncrétismes entre les radicaux à l'intégralité des formes verbales du premier binyan.

#### 5.3.3.1 Discussion et extension de l'analyse de Camilleri (2011)

L'analyse de Camilleri rend correctement compte de la grande majorité des données des radicaux des verbes du premier binyan en maltais. Néanmoins, la formalisation explicite et l'implémentation avec lexique associé ont mis en évidence un certain nombre de cas particuliers qui n'avaient pas été pris en compte dans l'analyse de Camilleri. Outre les omissions indiquées en bleu dans les tableaux ci-dessus, l'analyse ne captait notamment pas les cas de surabondance au sens de Thornton (2011) tels qu'illustrés par des verbes comme *SATAR* « protéger, couvrir ». *SATAR*, dont le paradigme est indiqué au tableau 5.19, présente en effet le comportement de deux CSR distincts au sens de Camilleri (2011) : la CSR  $ABCC\ AE$  d'une part et la CSR  $ABCC\ AE$  de l'autre.

L'analyse de Camilleri (2011) ne portant que sur les alternances de radicaux, elle ne permettait pas non plus de rendre compte des propriétés de la voyelle d'extension (ou voyelle thématique) apparaissant dans les formes imperfectives de certains verbes.

Le comportement de cette voyelle thématique est en général considéré comme résultant directement de la structure consonantique et de la couleur des voyelles des radicaux des verbes pour lesquels elle s'observe. La règle générale stipule ainsi :

1. La voyelle thématique de l'imperfectif apparaît uniquement pour les verbes dont la base verbale est de type CVCVC ou CVCV. Les autres bases verbales forment leur imperfectif sans voyelle thématique.

	SATAR « couvrir/protéger »		LAQAT « frapper »	BEŽAQ « cracher »		QALEB « changer »	
PFV 1SG	<i>stárt</i>	<i>stórt</i>	<i>lqátt</i>	<i>bžáqt</i>		<i>qlíbt</i>	
PFV 2SG	<i>stárt</i>	<i>stórt</i>	<i>lqátt</i>	<i>bžáqt</i>		<i>qlíbt</i>	
PFV 1PL	<i>stár.na</i>	<i>stór.na</i>	<i>lqát.na</i>	<i>bžáq.na</i>		<i>qlíb.na</i>	
PFV 2PL	<i>stár.tu</i>	<i>stór.tu</i>	<i>lqát.tu</i>	<i>bžáq.tu</i>		<i>qlíb.tu</i>	
PFV 3SG.M	<i>sá.tar</i>		<i>láqat</i>	<i>bé.žaq</i>		<i>qá.leb</i>	<i>qé.leb</i>
PFV 3SG.F	<i>sát.ret</i>		<i>láq.tet</i>	<i>béz.qet</i>		<i>qálb.et</i>	<i>qélb.et</i>
PFV 3PL	<i>sát.ru</i>		<i>láq.tu</i>	<i>béz.qu</i>		<i>qál.bu</i>	<i>qél.bu</i>
IPFV 1SG	<i>ní.stor</i>	<i>ní.star</i>	<i>nó.lqot</i>	<i>ní.bžoq</i>	<i>nó.bžoq</i>	<i>né.qleb</i>	<i>ná.qleb</i>
IPFV 2SG	<i>tí.stor</i>	<i>tí.star</i>	<i>tó.lqot</i>	<i>tí.bžoq</i>	<i>tó.bžoq</i>	<i>té.qleb</i>	<i>tá.qleb</i>
IPFV 3SG.M	<i>jí.stor</i>	<i>jí.star</i>	<i>jó.lqot</i>	<i>jí.bžoq</i>	<i>jó.bžoq</i>	<i>jé.qleb</i>	<i>já.qleb</i>
IPFV 3SG.F	<i>tí.stor</i>	<i>tí.star</i>	<i>tó.lqot</i>	<i>tí.bžoq</i>	<i>tó.bžoq</i>	<i>té.qleb</i>	<i>tá.qleb</i>
IPFV 1PL	<i>níst.ru</i>		<i>nólq.tu</i>	<i>níbž.qu</i>	<i>nóbž.qu</i>	<i>né.qil.bu</i>	<i>ná.qil.bu</i>
IPFV 2PL	<i>tíst.ru</i>		<i>tólq.tu</i>	<i>tíbž.qu</i>	<i>tóbž.qu</i>	<i>té.qil.bu</i>	<i>tá.qil.bu</i>
IPFV 3PL	<i>jíst.ru</i>		<i>jólq.tu</i>	<i>jíbž.qu</i>	<i>jóbž.qu</i>	<i>jé.qil.bu</i>	<i>já.qil.bu</i>
IMP 2SG	<i>ís.tor</i>	<i>ís.tar</i>	<i>ó.lqot</i>	<i>í.bžoq</i>	<i>ó.bžoq</i>	<i>é.qleb</i>	<i>á.qleb</i>
IMP 2PL	<i>íst.ru</i>		<i>ólq.tu</i>	<i>íbž.qu</i>	<i>óbž.qu</i>	<i>é.qil.bu</i>	<i>á.qil.bu</i>

TABLE 5.19 : Paradigmes surabondants en maltais

2. Pour les verbes ayant une voyelle thématique, cette dernière est en général la voyelle « *i* », cf. l'exemple standard de KITEB « *écrire* » au tableau 5.8 page 316.
3. Toutefois, si la première consonne du radical d'un verbe à base verbale CVCVC ou CVCV est une consonne arrière, la voyelle thématique de l'imperfectif est identique à la voyelle du radical de l'imperfectif (singulier). Un exemple de ce cas est le verbe HATAF « *saisir* » au tableau 5.8<sup>43</sup>.

Mais dans certains cas, les verbes ne se conforment pas à ces règles générales. C'est notamment le cas de LAQAT « *frapper* ». LAQAT est un verbe à base verbale CVCVC, il possède donc une voyelle thématique à l'imperfectif. Par ailleurs, la consonne initiale « *l* » de LAQAT n'est pas une consonne arrière. On s'attendrait donc à ce que la voyelle thématique affichée par LAQAT soit le « *i* » de la règle générale disant qu'en l'absence d'une première consonne de type arrière dans le radical, la voyelle thématique de l'imperfectif des verbes à base verbale CVCVC ou CVCV est par défaut la voyelle « *i* ». Or on observe dans le paradigme de LAQAT illustré au tableau 5.19 page 330 que la voyelle thématique

43. Les consonnes considérées comme des consonnes arrières en maltais correspondent aux graphèmes suivants : gh, h, h et q.

réalisée est « o ». Elle copie donc la voyelle de du radical de l'imperfectif singulier quand bien même la consonne initiale du verbe n'est pas une consonne arrière. Ce faisant elle contrevient à la règle générale qui semblait à première vue directement déductible de la phonologie du maltais, à savoir du type de consonne initiale d'un verbe donné. Ce type d'idiosyncrasies contredisant les règles générales nécessite donc un traitement au sein de la description morphologique.

Ce fait est d'autant plus marquant qu'il existe des verbes comme BEŻAQ « *cracher* » qui affichent même une surabondance en termes de voyelle thématique à l'imperfectif. Le paradigme de BEŻAQ est donné au tableau 5.19. Si plusieurs règles peuvent s'appliquer de façon concurrente à la même entrée il est en effet peu plausible que le résultat soit issu de la phonologie automatique.

Notons enfin que les deux phénomènes de surabondance (thématique et radicale) non couverts par l'analyse de Camilleri (2011) peuvent également se rencontrer en même temps, cf. le paradigme de QALEB « *changer* ». QALEB est surabondant au niveau radical aux troisièmes personnes du perfectif et surabondant au niveau thématique à l'imperfectif.

### 5.3.3.2 Analyse formelle et implémentation des verbes du premier binyan en maltais

L'analyse que nous avons conçue en  $\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}$  des verbes du premier binyan en maltais comporte trois niveaux réalisationnels :

- un niveau radical pour rendre compte des alternances radicales observées dès (Camilleri, 2011),
- un niveau thématique pour rendre compte de la présence de la voyelle thématique selon les verbes, ainsi que du type de voyelle présent le cas échéant,
- un niveau d'exponence simple comportant l'unique classe flexionnelle nécessaire.

Ces trois niveaux réalisationnels sont illustrés par le tableau 5.20 page 333 : le niveau radical en nuances de rouge, le niveau thématique en vert et le niveau d'exponence en bleu.

Le niveau radical comporte des zones de sélection de règles de formation de radical regroupables en cinq classes de radicaux, c'est à dire autant que de bases verbales dans l'analyse de Camilleri (2011). Les règles de réalisation associées aux flexèmes par ces zones de radicaux font usage d'un certain nombre d'opérations morphologiques définies comme indiqué à la figure 5.37 page 332 en Alexina  $\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}$ .

Les zones de radicaux sont ainsi définies pour chacun des espaces de syncrétisme systématique introduits dans l'analyse de Camilleri (2011). La figure 5.38 page 334 donne

```

1  <!-- Stem operations -->
2
3  <operation_definition name="deleteV1">
4  <replace source = "[[:C:]][0:V:][[:C:]][[:V:]][[:C:]]" target="[[:C:]][[:C:]][[:V
   :]][[:C:]]" rev="-" />
5  <replace source = "[[:C:]][0:V:][[:C:]][[:V:]]" target="[[:C:]][[:C:]][[:V:]]" rev=
   "-" />
6  </operation_definition>
7
8  <operation_definition name="deleteV2">
9  <replace source = "[[:C:]][[:V:]][[:C:]][0:V:][[:C:]]" target="[[:C:]][[:V:]][[:C
   :]][[:C:]]" rev="-" />
10 <replace source = "[[:C:]][[:V:]][[:C:]][[:C:]][0:V:][[:C:]]" target="[[:C:]][[:V
   :]][[:C:]][[:C:]]" rev="-" />
11 </operation_definition>
12
13 <operation_definition name="deleteV12">
14 <replace source = "[[:C:]][[:C:]][1:V:][[:C:]]" target="«[1:V:]»[:C:][[:C:]][[:C
   :]]" rev="-" />
15 <replace source = "[[:C:]][[:C:]][1:V:]" target="«[1:V:]»[:C:][[:C:]]" rev="-"
   />
16 </operation_definition>
17
18 <operation_definition name="changeV1">
19 <replace source = "[[:C:]][0:V:][[:C:]][[:V:]][[:C:]]" target="[[:C:]]_[:C:]][[:V
   :]][[:C:]]" rev="-" />
20 <replace source = "[[:C:]][0:V:][[:C:]][[:C:]][[:V:]][[:C:]]" target="[[:C:]]_[:C
   :]][[:C:]][[:V:]][[:C:]]" rev="-" />
21 <replace source = "[[:C:]][0:V:][[:C:]][[:C:]]" target="[[:C:]]_[:C:]][[:C:]]" rev
   ="- " />
22 <replace source = "[[:C:]][0:V:][[:C:]]" target="[[:C:]]_[:C:]]" rev="-" />
23 </operation_definition>
24
25 <operation_definition name="changeV2">
26 <replace source = "[[:C:]][[:V:]][[:C:]][[:C:]][0:V:][[:C:]]" target="[[:C:]][[:V
   :]][[:C:]][[:C:]]_[:C:]]" rev="-" />
27 </operation_definition>
28
29 <operation_definition name="deleteV1changeV2">
30 <replace source = "[[:C:]][0:V:][[:C:]][0:V:][[:C:]]" target="[[:C:]][[:C:]]_[:C
   :]]" rev="-" />
31 <replace source = "[[:C:]][0:V:][[:C:]][0:V:]" target="[[:C:]][[:C:]]_" rev="-"
   />
32 </operation_definition>
33
34 <operation_definition name="deleteV2changeV1">
35 <replace source = "[[:C:]][0:V:][[:C:]][0:V:][[:C:]]" target="[[:C:]]_[:C:]][[:C
   :]]" rev="-" />
36 </operation_definition>

```

FIGURE 5.37 : Opérations morphologiques sur les radicaux définies (données de MaltLex)

	Exp.	VE	RADICAL	Exp.
1SG			RAD2	-t
2SG			RAD2	-t
3SG.M			RAD1	
PFV	3SG.F		RAD3	-et
1PL			RAD2	-na
2PL			RAD2	-tu
3PL			RAD4	-u
IPFV	1SG	n- (v)	RAD5	
	2SG	t- (v)	RAD5	
	3SG.M	j- (v)	RAD5	
	3SG.F	t- (v)	RAD5	
	1PL	n- (v)	RAD6	-u
	2PL	t- (v)	RAD6	-u
	3PL	j- (v)	RAD6	-u

TABLE 5.20 : Les trois niveaux réalisationnels l'analyse  $\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}$  du maltais

pour exemple de zones les zones employées par les verbes à base verbale CVCVC. Notons que pour certaines variantes, des syncrétismes supplémentaires opèrent. Un exemple de cela est la règle de formation du RAD5 S5 pour les verbes à bases verbale CVCVC ayant la variante  $\text{var}=\text{AE}$  ou  $\text{var}=\text{AC}$ , c'est-à-dire correspondant respectivement aux CSR ABCC AE et ABCC AC dans la classification de Camilleri (2011).

Au niveau thématique, quatre zones sont définies : la zone des voyelles thématiques standard (figure 5.39 page 335) et les quatre zones spécifiques introduisant une voyelle particulière pour les verbes ne suivant pas la règle générale (figure 5.40 page 335).

La règle d'insertion de voyelle thématique standard insère un thème non spécifié  $\theta$  et s'accompagne par ailleurs d'une morphologie régulière garantissant que la voyelle thématique sera identique à celle du radical de l'imperfectif singulier pour les verbes à radical à consonne arrière initiale et « i » dans les autres cas. Ces règles morphologiques sont données à la figure 5.42 page 336. Elles reposent sur les classes de caractères définis en figure 5.41 page 335 et où B correspond aux consonnes arrières du maltais.

Par ailleurs, les verbes ne réalisant pas de voyelle thématique à l'imperfectif sont associés au niveau thématique avec la zone vide athématique  $\text{notheme}$  précisant l'absence de voyelle thématique, cf. la figure 5.43 page 336.

Enfin, tous les verbes sont associés à la même classe flexionnelle au niveau exposant. Cette classe est illustrée à la figure 5.44.



```

1  <!-- Stem level -->
2
3  <level type="stem" level="1">
4    <partitionspace name="S1" features="3.m.sg.pfv"/>
5    <partitionspace name="S2" features="1.pfv|2.pfv"/>
6    <partitionspace name="S3" features="3.f.sg.pfv"/>
7    <partitionspace name="S4" features="3.pl.pfv"/>
8    <partitionspace name="S5" features="sg.ipfv"/>
9    <partitionspace name="S6" features="pl.ipfv"/>
10
11
12  <!-- Stem tables -->
13
14  <table name="CVCVC" rads="[:C:][:V:][:C:][:V:][:C:]">
15    <item name="S1" source="S0" operation="changeV1(e)" var="2S1"/>
16    <item name="S1"/>
17
18    <item name="S2" source="S0" operation="deleteV1changeV2(o)" rads="
19      [:C:][:V:][:C:]a[:C:]" var="a2o"/>
20    <item name="S2" source="S1" operation="deleteV1changeV2(i)" rads="
21      [:C:][:V:][:C:]e[:C:]" />
22    <item name="S2" source="S1" operation="deleteV1" radsexcept="[:C:
23      :][:V:][:C:]e[:C:]" />
24
25    <item name="S3" source="S1" operation="deleteV2"/>
26
27    <item name="S4" source="S3" append=""/>
28
29    <item name="S5" source="S2" operation="" var="AE"/>
30    <item name="S5" source="S2" operation="" var="AC"/>
31    <item name="S5" source="S1" operation="deleteV1" rads="[:C:](e|i|a)
32      [:C:]e[:C:]" />
33    <item name="S5" source="S1" operation="deleteV1changeV2(i)" rads="
34      [:C:]a[:R:]a[:C:]" var="5Vi"/>
35    <item name="S5" source="S1" operation="deleteV1changeV2(o)" rads="
36      [:C:](a|e)[:C:]a[:C:]" />
37
38    <item name="S6" source="S3" operation="" var="DC"/>
39    <item name="S6" source="S3" operation="" var="AC"/>
40    <item name="S6" source="S1" operation="deleteV2changeV1(e)" rads="
41      [:C:]a[:R:]a[:C:]" var="AE"/>
42    <item name="S6" source="S1" operation="deleteV2changeV1(o)" rads="
43      [:C:]e[:R:]a[:C:]" var="DE"/>
44    <item name="S6" source="S1" operation="deleteV2changeV1(i)" rads="
45      [:C:]a[:R:]e[:C:]" />
46    <item name="S6" source="S1" operation="deleteV2changeV1(o)" rads="
47      [:C:]a[:R:][:V:][:C:]" />
48    <item name="S6" source="S1" operation="deleteV2changeV1(i)" rads="
49      [:C:]e[:R:][:V:][:C:]" />
50    <item name="S6" source="S5" operation="deleteV12" rads="[:C:][:nR
51      :][:V:][:C:]" />
52
53  </table>

```

FIGURE 5.38 : Règles de formation de radicaux pour les verbes à base verbale CVCVC  
334 (données de MaltLex)

```

1 <level type="theme" level="2">
2   <partitionspace name="PFV" features="pfv"/>
3   <partitionspace name="IPFV" features="ipfv"/>
4
5   <table name="theme" rads="..">
6     <item left_append="θ" features="ipfv" var="std"/>
7   </table>

```

FIGURE 5.39 : Règle de réalisation standard de la voyelle thématique CVCVC (données de MaltLex)

```

1 <table name="theme-a" rads="..">
2   <item left_append="a" features="ipfv" var="a"/>
3 </table>
4
5 <table name="theme-e" rads="..">
6   <item left_append="e" features="ipfv" var="e"/>
7 </table>
8
9 <table name="theme-i" rads="..">
10  <item left_append="i" features="ipfv" var="i"/>
11 </table>
12
13 <table name="theme-o" rads="..">
14  <item left_append="o" features="ipfv" var="o"/>
15 </table>

```

FIGURE 5.40 : Règles de réalisation spécifiques de la voyelle thématique CVCVC (données de MaltLex)

```

1 <!-- PHONOLOGY -->
2
3 <letterclass name="C" letters="b ċ d f ġ g ħ ħ h j k l m n p q r s t
4   v w x ž z ' "/>
5 <letterclass name="B" letters="għ ħ h q ' "/>
6 <letterclass name="V" letters="a e i o u ie"/>
7 <letterclass name="R" letters="għ l m n r"/>
8 <letterclass name="nR" letters="b ċ d f ġ g ħ h j k p q s t v w x ž z
9   ' "/>
10 <letterclass name="Vnonu" letters="a e i o ie"/>

```

FIGURE 5.41 : Classes de caractères définis pour le maltais (données de MaltLex)

```

1 <!-- MORPHONOLOGY -->
2 <!-- dealing with the stem extension vowel in std case-->
3
4
5 <sandhi source="θ_«a»[:B:]" target="a_[:B:]" final="+"/>
6 <sandhi source="θ_«e»[:B:]" target="e_[:B:]" final="+"/>
7 <sandhi source="θ_«o»[:B:]" target="o_[:B:]" final="+"/>
8 <sandhi source="«[0:V:]»" target="" final="+"/>
9
10 <sandhi source="θ_[:B:][:C:]a" target="a_[:B:][:C:]a" final="+"/>
11 <sandhi source="θ_[:B:][:C:]e" target="e_[:B:][:C:]e" final="+"/>
12 <sandhi source="θ_[:B:][:C:]o" target="o_[:B:][:C:]o" final="+"/>
13 <sandhi source="θ_" target="i_" final="+"/>

```

FIGURE 5.42 : Règles morphologiques liées à la voyelle thématique (données de MaltLex)

```

1 <table name="notheme" rads="..">
2 </table>

```

FIGURE 5.43 : Zone athématique (données de MaltLex)

```

1 <!-- Exponence level -->
2
3 <level type="exponent" level="3">
4
5 <table name="exponence" rads="">
6
7 <item block="1-2" suffix="na" features="1.pl.pfv"/>
8 <item block="1-2" suffix="et" features="3.f.sg.pfv"/>
9
10 <item block="1" suffix="t" features="1sg.pfv|2.pfv"/>
11
12 <item block="1" prefix="n" features="1.ipfv"/>
13 <item block="1" prefix="t" features="2.ipfv"/>
14 <item block="1" prefix="t" features="3.f.sg.ipfv"/>
15 <item block="1" prefix="j" features="3.ipfv"/>
16
17 <item block="2" suffix="u" features="pl"/>
18
19 </table>
20
21 </level>

```

FIGURE 5.44 : Classe flexionnelle unique des verbes maltais (données de MaltLex)

À partir de cette description, nous avons ensuite associé un schème à chaque entrée lexicale de notre lexique. Un échantillon des entrées intensionnelles pour les verbes KITEB « écrire », SATAR « protéger », LAQAT « frapper », BEŽAQ « cracher », et QALEB « changer » est donné à la figure 5.45 page 337.

```

1 kiteb v-CVCVC:DE:std
2 laqat v-CVCVC:DE:o
3 bezaq v-CVCVC:DE:std:o/: ,bzoq,bozq
4 satar v-CVCVC-S25over:AE:std
5 qaleb v-CVCVC-S134over:DE:a:e

```

FIGURE 5.45 : Entrées intensionnelles (données de MaltLex)

Les schèmes associés à chaque entrée ont été définis dans la grammaire. Nous donnons aux figures 5.46 page 340 à 5.48 page 341 les définitions des schèmes correspondant aux schèmes indiqués pour les verbes de la figure 5.45 page 337. Notons que certains sous-schèmes définis dans ces schèmes dépendent des variantes définies dans les entrées lexicales, cf. la variante `var="o"` pour les verbes ayant une voyelle thématique non standard « o » comme LAQAT « frapper ». Ainsi, les entrées compatibles avec plusieurs sous-schèmes comme BEŽAQ pourront utiliser tous ceux avec lesquels ils sont compatibles et seront donc surabondants. En plus du sous-schème définissant la flexion au perfectif<sup>44</sup>, BEŽAQ se fléchira ainsi à l'imperfectif<sup>45</sup> selon le sous-schème pour les verbes à voyelle thématique « e » (`var=e`) et le sous-schème pour les verbes à voyelle thématique « o » irrégulière (`var=o`). KITEB « écrire » n'emploiera que le sous-schème standard à l'imperfectif et LAQAT « frapper » que le sous-schème irrégulier en `var=o`.

Les verbes SATAR « protéger » et QALEB « changer » sont par ailleurs surabondants au niveau radical. Ils ont donc leurs schèmes spécifiques illustrés respectivement par les figures 5.47 page 341 et 5.48 page 341.

Dans le cas de SATAR nous observons une surabondance pour les formes de la première et deuxième personne du perfectif ainsi qu'à l'imperfectif singulier. En associant ces formes à un espace partitionnant noté CH-a<sub>o</sub>, nous pouvons poser un sous-schème permettant de réaliser l'intégralité des formes de la première et deuxième personnes du perfectif et de l'imperfectif singuliers avec un RAD2 pour lequel la seconde voyelle subit un changement vocalique  $a \rightarrow o$ . La règle permettant ce changement vocalique est associée à la variante

44. Sous-schème indiquant pour le niveau réalisationnel thématique `level="2"` `partitionspace="PFV"`.

45. Sous-schèmes indiquant pour le niveau réalisationnel thématique `level="2"` `partitionspace="IPFV"`.

$\text{var} = \text{a2o}$  dans l'implémentation. Cette variante est également précisée dans le premier sous-schème. *SATAR* étant un verbe à syncrétisme entre le *RAD2* et le *RAD5* ( $\text{var} = \text{AE}$ ), le *RAD5* est obtenu directement à partir du *RAD2*. Par ailleurs, un second sous-schème produit l'intégralité des formes sans changement vocalique au *RAD2*.

Pour *QALEB*, un sous-schème spécifique défini pour un espace partitionnant  $3\text{PFV}$  recouvrant les seules formes de la troisième personne du perfectif permet de construire les formes avec le radical surabondant comportant un changement vocalique  $a \rightarrow e$  sur la première voyelle. Ce sous-schème est associé à la variable  $2S1$  dans l'entrée lexicale. Les autres radicaux sont contruits de façon régulière à l'aide des autres sous-schèmes. Par ailleurs, la surabondance thématique est prise en compte par la compatibilité de *QALEB* avec les deux variantes non standard *a* et *e*.

En figure 5.49, nous indiquons en conclusion les entrées extensionnelles produites par notre description pour l'entrée lexicale *QALEB*.

### 5.3.3.3 Les apports de l'implémentation pour l'analyse de la flexion maltaise

En conclusion de cette analyse nous remarquons que la formalisation nous a forcée à prendre en compte l'intégralité de la réalisation des formes et a ainsi mis l'accent sur la complexité inhérente au choix de la voyelle thématique de l'imperfectif en maltais. L'implémentation avec lexique associé a par ailleurs révélé des cas de surabondance qui n'avaient pas été notés par les analyses antérieures<sup>46</sup> ainsi qu'un certain nombre de changements vocaliques omis.

Notre description couvre ainsi désormais les différents types de surabondance que l'on rencontre dans les paradigmes verbaux du premier binyan en maltais. La cohérence de l'analyse a pu être vérifiée par la génération par les outils de flexion liés à *Alexina<sub>PARSLI</sub>* de près de 10 000 formes des quelque 600 verbes de notre lexique *MaltLex* et leur validation partielle par une locutrice native de la langue.

## 5.4 Analyses implémentées et choix descriptifs

Dans ce chapitre, nous avons présenté le formalisme lexical *Alexina<sub>PARSLI</sub>* permettant d'implémenter des analyses *PARSLI*. En tant que formalisme lexical il permet d'associer toute description morphologique à un lexique à grande échelle permettant de vérifier non seulement la cohérence d'une analyse proposée, mais également sa couverture par rapport

---

46. La grammaire correspondant à l'implémentation complète et son lexique associé sont disponibles sur la page personnelle de l'auteur.

à la quasi-exhaustivité du système flexionnel d'une langue donnée. En revanche, le fait de pouvoir implémenter une description garantit seulement sa cohérence et sa couverture. Cela ne signifie pas pour autant que la description donnée est la seule cohérente ou la meilleure description possible. Nous avons évoqué plus haut le fait que de nombreuses descriptions de la flexion verbale des langues sémitiques posent l'existence de classes flexionnelles concurrentes pour rendre compte des différents paradigmes. La segmentation proposée dans ces analyses diffère ainsi de celle proposée par Camilleri (2011).

Il serait bien évidemment possible de concevoir et d'implémenter une description  $\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}$  de la flexion verbale maltaise qui suivrait cette tradition. Correctement posée et aussi exhaustivement implémentée, elle ne sera ni plus ni moins cohérente ou couvrante que l'analyse que nous avons proposée à la section 5.3.3. Si l'on veut malgré tout pouvoir comparer entre elles des analyses concurrentes, au préalable vérifiées par une implémentation à grande échelle quant à leur cohérence et leur couverture, il sera nécessaire d'avoir accès à un critère d'évaluation objectif externe. C'est la mise en place d'un tel critère qui fera l'objet du chapitre 6.

```

1 <pattern name="v-CVCVC" cat="v">
2   <subpattern >
3     <realzone level="1" table="CVCVC"/>
4     <realzone level="2" partitionospace="PFV" table="theme"/>
5     <realzone level="3" table="exponence"/>
6   </subpattern>
7
8   <subpattern>
9     <realzone level="1" table="CVCVC"/>
10    <realzone level="2" partitionospace="IPFV" table="theme" var="std"/>
11    <realzone level="3" table="exponence"/>
12  </subpattern>
13  <subpattern>
14    <realzone level="1" table="CVCVC"/>
15    <realzone level="2" partitionospace="IPFV" table="theme-a" var="a"/>
16    <realzone level="3" table="exponence"/>
17  </subpattern>
18  <subpattern>
19    <realzone level="1" table="CVCVC"/>
20    <realzone level="2" partitionospace="IPFV" table="theme-e" var="e"/>
21    <realzone level="3" table="exponence"/>
22  </subpattern>
23  <subpattern>
24    <realzone level="1" table="CVCVC"/>
25    <realzone level="2" partitionospace="IPFV" table="theme-i" var="i"/>
26    <realzone level="3" table="exponence"/>
27  </subpattern>
28  <subpattern>
29    <realzone level="1" table="CVCVC"/>
30    <realzone level="2" partitionospace="IPFV" table="theme-o" var="o"/>
31    <realzone level="3" table="exponence"/>
32  </subpattern>
33 </pattern>

```

FIGURE 5.46 : Schème des verbes non surabondants (données de MaltLex)

```

1 <pattern name="v-CVCVC-S25over" cat="v">
2   <subpattern >
3     <realzone level="1" partitionspace="S2+S5" table="CVCVC" local_var=
4       "a2o"/>
5     <realzone level="2" table="theme"/>
6     <realzone level="3" table="exponence"/>
7   </subpattern>
8
9   <subpattern >
10    <realzone level="1" table="CVCVC" />
11    <realzone level="2" table="theme"/>
12    <realzone level="3" table="exponence"/>
13  </subpattern>
14 </pattern>

```

FIGURE 5.47 : Schème de SATAR « protéger » (données de MaltLex)

```

1 <pattern name="v-CVCVC-S134over" cat="v">
2   <subpattern >
3     <realzone level="1" partitionspace="S1+S3+S4" table="CVCVC"
4       local_var="2S1"/>
5     <realzone level="2" partitionspace="PFV" table="theme"/>
6     <realzone level="3" table="exponence"/>
7   </subpattern>
8
9   <subpattern >
10    <realzone level="1" table="CVCVC"/>
11    <realzone level="2" partitionspace="PFV" table="theme"/>
12    <realzone level="3" table="exponence"/>
13  </subpattern>
14
15  <subpattern>
16    <realzone level="1" table="CVCVC"/>
17    <realzone level="2" partitionspace="IPFV" table="theme-a" local_var
18      ="a"/>
19    <realzone level="3" table="exponence"/>
20  </subpattern>
21
22  <subpattern>
23    <realzone level="1" table="CVCVC"/>
24    <realzone level="2" partitionspace="IPFV" table="theme-e" local_var
25      ="e"/>
26    <realzone level="3" table="exponence"/>
27  </subpattern>
28 </pattern>

```

FIGURE 5.48 : Schème de QALEB « changer » (données de MaltLex)



```

1 qaleb qlibt unmarked.pfv.sg.1 v-CVCVC-S134over:DE:a:e
2 qaleb qlibt unmarked.pfv.sg.2 v-CVCVC-S134over:DE:a:e
3 qaleb qaleb unmarked.pfv.sg.3.m v-CVCVC-S134over:DE:a:e
4 qaleb qaleb unmarked.pfv.sg.3.m v-CVCVC-S134over:DE:a:e
5 qaleb qalbet unmarked.pfv.sg.3.f v-CVCVC-S134over:DE:a:e
6 qaleb qelbet unmarked.pfv.sg.3.f v-CVCVC-S134over:DE:a:e
7
8 qaleb qlibna unmarked.pfv.pl.1 v-CVCVC-S134over:DE:a:e
9 qaleb qlibtu unmarked.pfv.pl.2 v-CVCVC-S134over:DE:a:e
10 qaleb qalbu unmarked.pfv.pl.3 v-CVCVC-S134over:DE:a:e
11 qaleb qelbu unmarked.pfv.pl.3 v-CVCVC-S134over:DE:a:e
12
13
14 qaleb naqleb unmarked.ipfv.sg.1 v-CVCVC-S134over:DE:a:e
15 qaleb neqleb unmarked.ipfv.sg.1 v-CVCVC-S134over:DE:a:e
16 qaleb taqleb unmarked.ipfv.sg.2 v-CVCVC-S134over:DE:a:e
17 qaleb teqleb unmarked.ipfv.sg.2 v-CVCVC-S134over:DE:a:e
18 qaleb jaqleb unmarked.ipfv.sg.3.m v-CVCVC-S134over:DE:a:e
19 qaleb jeqleb unmarked.ipfv.sg.3.m v-CVCVC-S134over:DE:a:e
20 qaleb taqleb unmarked.ipfv.sg.3.f v-CVCVC-S134over:DE:a:e
21 qaleb teqleb unmarked.ipfv.sg.3.f v-CVCVC-S134over:DE:a:e
22
23 qaleb naqilbu unmarked.ipfv.pl.1 v-CVCVC-S134over:DE:a:e
24 qaleb neqilbu unmarked.ipfv.pl.1 v-CVCVC-S134over:DE:a:e
25 qaleb taqilbu unmarked.ipfv.pl.2 v-CVCVC-S134over:DE:a:e
26 qaleb teqilbu unmarked.ipfv.pl.2 v-CVCVC-S134over:DE:a:e
27 qaleb jaqilbu unmarked.ipfv.pl.3 v-CVCVC-S134over:DE:a:e
28 qaleb jeqilbu unmarked.ipfv.pl.3 v-CVCVC-S134over:DE:a:e
29
30 qaleb aqleb imp.ipfv.sg v-CVCVC-S134over:DE:a:e
31 qaleb eqleb imp.ipfv.sg v-CVCVC-S134over:DE:a:e
32 qaleb aqilbu imp.ipfv.pl v-CVCVC-S134over:DE:a:e
33 qaleb eqilbu imp.ipfv.pl v-CVCVC-S134over:DE:a:e

```

FIGURE 5.49 : Entrées extensionnelles de QALEB « *changer* » (données de MaltLex)

# Évaluations externes de descriptions morphologiques

Frustra fit per plura quod potest fieri  
per pauciora.

---

Guillaume d'Okham (1323)

*Summa Totius Logica*

Aux chapitres 4 et 5 de cette partie, nous avons respectivement introduit le modèle formel de la morphologie flexionnelle  $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$  et son implémentation Alexina $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$ . Dans ce chapitre, nous introduisons une mesure externe permettant d'évaluer la qualité d'une description morphologique flexionnelle donnée.

Au sein d'un modèle de la morphologie flexionnelle comme le nôtre, qui représente la construction des formes comme l'application successive de règles réalisationnelles, un des problèmes principaux rencontrés est celui de la segmentation optimale entre les différents niveaux réalisationnels. Cette question surgit de façon récurrente dans toutes les approches tentant de proposer une segmentation de formes, notamment en radical et exposants associés. Cette segmentation est rarement indiscutable. En général, plusieurs analyses concurrentes, utilisant des segmentations différentes, peuvent rendre compte de mêmes données. Un exemple de cela est la représentation de la flexion verbale en français. La représentation traditionnelle du Bescherelle (Arrivé, 1997) propose une analyse de la flexion verbale du français utilisant 88 tables de flexion (ou classes flexionnelles) différentes, misant sur une complexité relativement élevée au niveau exposant. L'analyse de Bonami & Boyé, notamment Bonami et al. (2008) et Bonami & Boyé (2010a), en revanche, ne pose qu'une

---

classe flexionnelle unique et relaie la complexité de la flexion dans les alternances radicales représentées par les espaces thématiques comportant jusqu'à douze cases distinctes selon les verbes. Choisir la meilleure solution entre de telles analyses concurrentes n'est pas toujours immédiatement possible, au point que souvent, le choix est plus une question de goût que la conséquence de véritables critères objectifs.

Cette difficulté à pouvoir choisir objectivement une meilleure description, comportant notamment une segmentation idéale, peut mener à un rejet complet des approches CONSTRUCTIVES au profit d'analyses ABSTRACTIVES comme la MORPHOLOGIE IMPLICATIVE au sens de Blevins (2006). Ces dernières reposent sur des principes d'analogie entre cases de paradigmes plutôt que sur la recherche de règles de réalisation partielles permettant de décrire la construction des formes segmentées à partir de généralisations partielles sur des radicaux ou des exposants définis au préalable (Blevins, 2012). La question première dans l'étude de la morphologie flexionnelle devient alors celle du degré d'incertitude sur le mode de remplissage des cases des paradigmes, le *Paradigm Cell Filling Problem (PCFP)* (Ackerman et al., 2009)<sup>1</sup>, et non plus la spécification idéale des généralisations formulables sur des règles de réalisation partielles. Parmi ces approches, certaines, plus récentes utilisent l'*entropie conditionnelle* pour mesurer le degré d'incertitude.

Comme le note à juste titre Blevins (2006, 2012), l'arbitraire des découpages dans les démarches constructives constitue un défaut majeur de ces dernières. Toutefois, les généralisations qu'elles cherchent à faire sur les éléments de construction des formes, notamment le repérage de classes flexionnelles et l'identification des phénomènes non canoniques principaux que nous avons présentés au chapitre 3, nous donnent des informations sur la structure des paradigmes qui sont complémentaires de la question du remplissage des cases paradigmatiques traitées par la morphologie implicative. Rejeter purement et simplement les approches constructives aurait ainsi le regrettable effet de nous priver de ces informations auxquelles la morphologie implicative n'a — du moins à ce stade de son évolution — pas accès<sup>2</sup>. Par ailleurs, les travaux déjà développés au sein du paradigme constructif, notamment en typologie générale, peuvent facilement être exploités si l'on reste dans ce même paradigme, mais ils ne sont pas d'une grande utilité si l'on se cantonne au seul paradigme abstraitif. Nous adoptons ici plutôt une démarche inverse, à savoir **accepter la part d'arbitraire que comportent intrinsèquement les approches**

---

1. Il est toutefois à noter que Ackerman et al. (2009, p. 66 ss.) utilisent en fin de compte implicitement les notions de radical et exposant pour rendre compte des données qu'ils analysent.

2. Il est possible qu'à terme des structures internes puissent être induites directement par les approches abstraitives, mais cela reviendrait à induire automatiquement une description morphologique, ce qu'à ce jour les outils disponibles ne permettent pas encore. Par ailleurs, la question de la représentation de cette structure redeviendrait à ce moment de nouveau une question pertinente.

constructives et chercher un moyen de la contenir par des évaluations objectives externes sur les descriptions concurrentes d'un même phénomène — par exemple sur la modélisation de la flexion verbale selon Bonami & Boyé (2010a) et selon le Bescherelle (Arrivé, 1997). Dans ce chapitre nous proposons ainsi une mesure externe permettant d'évaluer quantitativement deux descriptions concurrentes d'un même (sous-)système flexionnel et de trancher en faveur de la *meilleure* d'entre elles selon un critère objectif déterminé.

Notons en un premier temps, comme nous l'avons argumenté à la section 5.1 du chapitre 5, que le premier critère d'évaluation d'une description de la flexion d'une (sous-partie) d'une langue est la cohérence et la complétude de la description. Ces deux critères peuvent aisément être vérifiés par une description implémentée (vérifiant l'exactitude des règles flexionnelles définies) et un lexique étendu associé (vérifiant la couverture de la description par rapport à l'inventaire (quasi-)exhaustif des phénomènes observés).

Mais au delà de ces exigences minimales, les descriptions de la flexion d'une langue peuvent encore grandement différer. En particulier, certaines descriptions ne distingueront pas suffisamment entre des phénomènes marquants, mais ponctuels, et d'autres dont la présence est plus massive dans le système flexionnel global. De telles descriptions ne parviendront ainsi pas à capter l'importance des phénomènes décrits par rapport au système flexionnel global.

Enfin, comme l'indique d'ailleurs Blevins (2006, 2012), l'idée principale derrière les approches constructives est de rendre compte de façon compacte d'un système flexionnel donné. En cela, une analyse constructive pourra être considérée comme d'autant meilleure qu'elle réussira à produire une description plus compacte du système analysé. Ces approches reposent sur l'idée ancienne sous-tendant toute démarche scientifique à savoir celle du rasoir d'Okham. Le principe du rasoir d'Okham consiste à dire que *toutes choses étant égales par ailleurs* (en l'occurrence la cohérence et la couverture d'une description étant assurées) *une description plus simple sera toujours préférable à une description plus complexe*. Autrement dit, une description d'un système sera d'autant meilleure qu'elle reflètera simplement la richesse dudit système.

Lors de la mise en place d'une description  $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$ , un certain nombre d'arbitrages sont cependant à faire directement par le descripteur. Ainsi, il aura par exemple à choisir le nombre de niveaux réalisationnels utilisés pour rendre compte du système qu'il décrit. Au chapitre 1, nous avons ainsi présenté une description du système verbal du latin à trois niveaux. Au chapitre 7, nous en proposerons une autre, à quatre niveaux. Choisir entre ces deux propositions, on le concède, arbitraires requiert l'existence de mesures

---

objectives externes de la compacité des systèmes proposés. Afin de mesurer la compacité des descriptions  $\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}$ , nous avons donc mis en place des mesures de descriptions s'appuyant sur la notion de LONGUEUR DE DESCRIPTION MINIMALE (LDM) — *Minimum description Length (MDL)* (Rissanen, 1984) — empruntée à la théorie de l'information. Cette mesure, décrite à la section 6.1 ci-dessous, permet ainsi d'évaluer la simplicité d'une description donnée au sens d'une mesure de compacité globale de la description proposée. Optimiser la compacité globale d'une description donnée peut ainsi servir de critère d'évaluation global d'un échantillon de descriptions concurrentes.

Nous avons ainsi obtenu une mesure permettant de trancher, par la comparaison quantitative de la compacité de deux descriptions  $\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}$  concurrentes, des questions descriptives fondamentales comme celles indiquées ci-dessous. Dans la mesure où chacune de ces décisions est susceptible d'affecter la compacité globale, nous pourrions ainsi décider

- si la segmentation en  $n$  niveaux d'une description A est préférable à la segmentation en  $m$  niveaux d'une description B
- si la répartition de l'information entre les niveaux (comme la segmentation entre radical et exposants) d'une description A est préférable à celle d'une description B
- si poser l'existence d'une classe flexionnelle supplémentaire est préférable à celle de poser l'existence d'un schème hétéroclite pour certaines entrées, cf. les données du slovaque décrites dans les chapitres 3 à 5
- si poser l'existence d'une classe flexionnelle à exponence homonymique supplémentaire est préférable à la celle de poser l'existence de cas de décalage morphosyntaxiques comme proposé pour les noms croates dans les données décrites aux chapitres 3 à 5
- si poser l'existence d'une catégorie flexionnelle supplémentaire est préférable à poser l'existence d'un certain nombre de flexèmes défectifs ou surabondants dans une des autres catégories existantes

*etc.*

Si nous employons notre mesure d'évaluation externe de la compacité, le choix revient ainsi toujours simplement à prendre l'option menant à une description plus compacte au sens de notre métrique. Certes, il est possible, entre deux descriptions concurrentes, que la différence en termes de compacité soit minime. Mais pouvoir affirmer quantitativement l'absence d'une différence significative entre deux descriptions concurrentes nous semble également constituer une information intéressante en tant que telle. Nous n'affirmons donc pas qu'il est toujours possible de décider qu'une description est *meilleure* qu'une autre en termes de compacité, ni qu'il est à possible, à ce stade, de trouver la description la plus

compacte d'entre toutes. Mais nous donnons ici une mesure permettant de déterminer si oui ou non une description donnée est plus compacte que sa concurrente, et ce de façon non négligeable ou non.

Le fait de vouloir évaluer la pertinence d'une analyse par rapport à une autre en termes d'économie descriptive n'est par ailleurs pas une idée nouvelle. Dans Jackendoff (1975), l'auteur introduit sa théorie du lexique et de l'interface morphologie-lexique en prenant soin d'accompagner son modèle d'un moyen d'évaluation du ratio *information-lexicale originale/règle de redondance* nécessaire pour introduire ou apprendre un mot qui ne se trouvait pas dans le lexique. La mesure de Jackendoff est ainsi à sa manière un travail précurseur de l'approche consistant à évaluer la longueur de description minimale que nous proposons dans ce chapitre.

Enfin, nous montrons dans ce chapitre que les questions de compacité de descriptions ont également des conséquences sur les questions de répartition de l'information entre les composantes de la grammaire d'une langue. Dans une vision modulaire intégrée de la morphologie flexionnelle comme celle que nous avons présentée en conclusion du chapitre 2 et que nous réillustrons ici à la figure 6.1, les questions de la répartition entre la morphologie flexionnelle d'une part et d'autre part les autres composantes, notamment le lexique, la morphonologie et la syntaxe, s'apparentent en fin de compte aux questions de segmentation au sein de la morphologie flexionnelle. L'arbitrage de la répartition entre les différents modules peut en effet se faire selon le même critère de compacité globale de la description. La meilleure analyse (au sens du rasoir d'Okham) d'un système linguistique sera alors celle qui maximisera la compacité globale de la description de ce dernier. Nous développons ce point en détail dans les sections 6.2 à 6.4 de ce chapitre.

Notons enfin que la mesure de longueur de description que nous mettons en place se veut être une mesure d'évaluation de descriptions concurrentes pour une langue donnée, voire pour un sous-système de cette langue. Elle ne cherche pas à servir de mesure de comparaison de la complexité de langues comme ont pu chercher à faire, à travers d'autres types de mesures, un certain nombre de travaux récents, voir par exemple (McWhorter, 2001; Juola, 2008; Bane, 2008; Goldsmith, 2001; Xanthos, 2008; Brown & Evans, 2012)<sup>3</sup>. Dans l'approche adoptée dans cette thèse, la comparaison entre systèmes est laissé

---

3. Ces dernières années, le développement de méthodes de mesure de la complexité linguistique est devenu un domaine de recherche actif. Par ordre décroissant de généralité, des travaux ont été réalisés dans ce domaine sur les langues prises de façon globale (McWhorter, 2001; Juola, 2008), en restreignant le travail à un niveau d'analyse particulier tel que la morphologie (Bane, 2008) et en mesurant la complexité de descriptions morphologiques particulières, notamment dans le contexte de l'acquisition automatique non supervisée ou faiblement supervisée de la morphologie (Goldsmith, 2001; Xanthos, 2008). Différentes métriques pour mesurer la complexité linguistique sont décrites dans la littérature. Les plus simples se contentent de compter

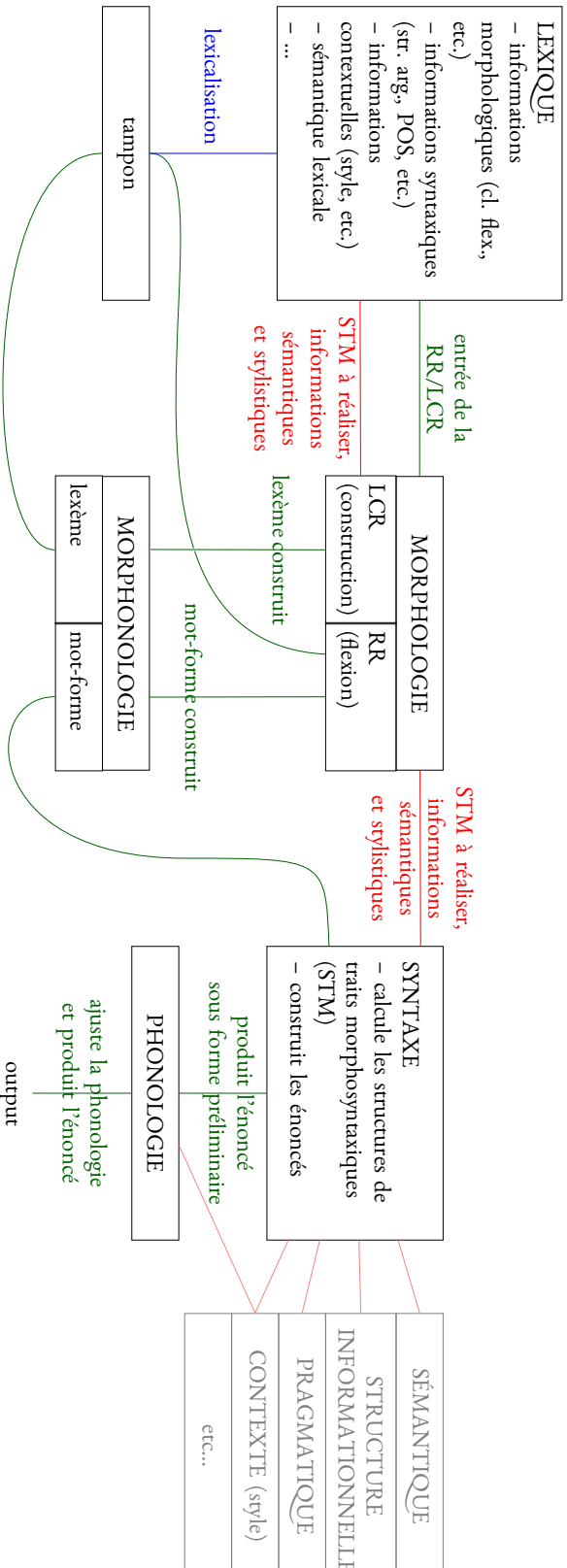


FIGURE 6.1 : Esquisse de l'architecture de la grammaire

à des mesures de canonicité telles que celles mises en place au chapitre 4<sup>4</sup>. Les mesures de compacité ne sont employées que pour évaluer au préalable les descriptions sur lesquelles les mesures de canonicité seront appliquées. Il paraît en effet essentiel de pouvoir donner une évaluation d'une description avant d'effectuer sur elle des mesures destinées à terme à servir de point de comparaison entre des systèmes de langues différentes.

Ce chapitre est organisé de la façon suivante. Nous présentons en un premier temps, à la section 6.1, la notion de longueur de description minimale (LDM) et montrons sa pertinence comme mesure d'évaluation de description de systèmes morphologiques flexionnels. Nous y détaillons en particulier la méthode que nous avons mise en place pour mesurer des descriptions  $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}\mathcal{I}$  implémentées en *Alexina $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}\mathcal{I}$*  et notamment la distribution de l'information entre les différentes composantes s'interfaçant avec la composante flexionnelle de la langue. À la section 6.2, nous montrons une application de notre mesure à quatre descriptions de la flexion verbale du français. Les sections 6.2 à 6.4 illustrent à l'aide de trois exemples respectivement du français, du maltais et du kurde sorani l'impact que peut avoir sur la compacité d'une description donnée la répartition de l'information relative à la flexion entre la grammaire flexionnelle d'une part et respectivement le lexique flexionnel, la composante (mor-)phonologique et la composante syntaxique de l'autre. Nous terminons ce chapitre par un récapitulatif sur l'importance de la distribution optimale de l'information entre les composantes de la grammaire et les mécanismes de factorisation.

## 6.1 Mesurer la longueur de description

Dans cette section nous présentons la mesure que nous avons retenue pour l'évaluation des descriptions réalisées dans notre modèle formel de la morphologie flexionnelle  $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}\mathcal{I}$ . La mesure en question est une mesure rendant compte de la compacité d'une description donnée. Son application sera illustrée à la section 6.2. Le but est de pouvoir évaluer de façon externe la description proposée avant d'appliquer à cette description les mesures de

---

les occurrences d'un ensemble de propriétés linguistiques définies manuellement : taille de divers inventaires (par exemple, inventaires des phonèmes, des catégories, des types de morph(èm)es...) (McWhorter, 2001). Cependant, de telles approches sont intrinsèquement biaisées : tant l'ensemble de propriétés choisies que les critères selon lesquels ces propriétés sont décrites sont très difficiles à définir selon des principes clairs, objectifs et reproductibles (par exemple, comment construire d'une façon objective et indépendante de la langue un inventaire de parties du discours ?). Des approches alternatives de mesure de la complexité existent, dont la plupart reposent sur des définitions de la complexité qui sont issues de la THÉORIE DE L'INFORMATION.

4. Ces mesures sont elles aussi complémentaires de celles qui ont pu être proposées par d'autres auteurs pour qualifier à chaque fois des caractéristiques différentes de systèmes flexionnels (voir par exemple le PCFP d'Ackerman et al. (2009)).



canonicité définies au chapitre 4. Dans les sections 6.2 à 6.4, nous montrerons notamment l'impact de la répartition de l'information entre les composantes de la langue pour la compacité globale de la description d'une langue.

### 6.1.1 La complexité de Kolmogoroff

Notre mesure d'évaluation externe de descriptions morphologiques flexionnelles repose sur une définition de la complexité issue de la théorie de l'information. Dans ce cadre, deux définitions différentes ont été utilisées dans les travaux récents, qui s'appliquent toutes deux à n'importe quel type de message et pas seulement aux descriptions ou aux modèles linguistiques : l'*entropie informationnelle* (ou *complexité de Shannon*<sup>5</sup>), et l'*entropie algorithmique* (ou *complexité de Kolmogoroff*), un moyen plus générique et plus objectif de mesure de la quantité d'informations contenue dans un message, mais qui n'est pas directement calculable et pour lequel on doit donc se contenter d'approximations.

La complexité de Kolmogoroff, parce qu'elle est plus générale et ne nécessite pas de modèle probabiliste sous-jacent, est plus adaptée à nos besoins. Elle repose sur l'intuition suivante : un modèle est plus complexe qu'un autre s'il nécessite un message plus long pour être décrit. Cependant, puisque le calcul de cette complexité est impossible directement, ce calcul est souvent ramené à celui d'un certain type d'entropie au sein d'un espace particulier de descriptions possibles, au moyen d'une approximation de la complexité de Kolmogoroff définie sur cet espace : le résultat est ce que l'on appelle la *longueur de description* par rapport au modèle. Cette notion est le fondement du paradigme d'apprentissage automatique appelé *Longueur de Description Minimale* (*Minimum Description Length*, ou *MDL*) développé par Rissanen (1984). Dans ce paradigme, l'apprentissage automatique se fait *via* l'identification de la description que l'on peut faire des données d'apprentissage qui ait une *longueur de description* minimale par rapport à un modèle choisi. Pour que la longueur de description soit une bonne approximation de la complexité de Kolmogoroff, il faut toutefois que l'on se dote d'une façon aussi efficace que possible de représenter une description, ici une description linguistique, sous la forme d'une chaîne de caractères (le code) (Bane, 2008). De plus, un modèle linguistique est le plus souvent structuré<sup>6</sup>, ce qui n'est pas toujours exploité par les travaux faisant usage de la complexité morphologique. En

---

5. Son principal inconvénient est qu'il nécessite l'encodage du message sous forme de séquence de variables aléatoires indépendantes et identiquement distribuées selon un certain modèle probabiliste, ce qui est difficile en pratique.

6. Certaines sections décrivent les opérations morphologiques existantes, d'autres le fonctionnement de l'allomorphie, d'autres encore les classes flexionnelles elles-mêmes structurées par les formes qu'elles génèrent *etc.*

particulier, l'évaluation de la complexité d'une représentation du lexique morphologique ne peut se réduire à mesurer la complexité d'un corpus dont les formes ont été segmentées en morphes — approche qui est toutefois le fondement de travaux pionniers en acquisition automatique d'informations morphologiques Goldsmith (2001).

### 6.1.2 La longueur de description minimale (LDM)

Dans notre cas, nous souhaitons mesurer la compacité de descriptions (d'une partie) de la morphologie d'une langue donnée dans le modèle  $\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}$ . Pour cela, nous avons développé une métrique à partir du paradigme d'apprentissage automatique Longueur de description Minimale (LDM) développé par Rissanen (1984).

La longueur de description  $DL(m)$  d'un message non structuré  $m$  au sein d'un modèle donné qui le code sous la forme d'une séquence de  $N$  symboles d'un alphabet  $W = \{w_1, \dots, w_n\}$  est définie par :

$$DL(m) = - \sum_i o(w_i) \log_2(o(w_i)/N)$$

où  $o(w_i)$  est le nombre d'occurrences de  $w_i$  dans  $m$ . Cette longueur de description est égale à  $N$  fois l'entropie du message.

Dans notre cas, il est pertinent d'utiliser un codage plus complexe qui rende compte du caractère structuré des descriptions morphologiques  $\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}$ . Tout d'abord, elle se décompose en un lexique et une grammaire flexionnels. En Alexina $\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}$ , une entrée lexicale se définit comme suit :

- une forme canonique
- un schème
- une liste optionnelle de radicaux non prédictibles (les radicaux réguliers sont pris en compte par la grammaire)
- des variantes optionnelles
- des classes de défektivité

Quant à la grammaire morphologique, elle met en jeu les composantes suivantes

- une liste de segments minimaux (phonèmes ou graphèmes)
- des règles de morphologiques ou morphographémiques
- un mécanisme de définition de traits pour calculer le nombre de cases à réaliser pour chaque catégorie.

- des procédés de construction de formes (affixation et autres opérations plus complexes...)
- des schèmes et sous-schèmes
- des zones réalisationnelles
- des classes flexionnelles, thématiques et de radicaux optionnelles,

Nous avons défini un codage qui représente toute cette structure de façon bijective (il peut être décodé de façon non ambiguë) au moyen de symboles de 13 alphabets différents<sup>7</sup>. Les alphabets que nous avons définis correspondent aux éléments suivants d'une description flexionnelle en Alexina<sup>PATRBLI</sup> :

1. un alphabet `patternchar` pour rendre compte respectivement des schèmes spécifiés pour chaque entrée lexicale dans le lexique et pour leur définition dans la grammaire ;
2. un alphabet `variantchar` pour rendre compte respectivement des variantes spécifiées pour chaque entrée lexicale dans le lexique et pour leur définition dans la grammaire ; cet alphabet comporte également un symbole pour l'absence de variante quand elle n'est pas explicitement indiquée dans la description ;
3. un alphabet `canformchar` pour rendre compte des formes de citation listées dans le lexique ;
4. un alphabet `supplchar` pour rendre compte des radicaux supplétifs et des formes supplétives dans le lexique ; cet alphabet couvre aussi, si nécessaire, les formes supplétives spécifiées manuellement dans le fichier contenant les formes supplémentaires ;
5. un alphabet `sandhi char` pour rendre compte des règles morphographémiques ou morphologiques ;
6. un alphabet `letterclasschar` pour rendre compte des segments minimaux ;
7. un alphabet `model` pour rendre compte de la structure générale de la description (début et fin de zones, de schèmes, de classes flexionnelles, de blocs *etc.*) ;
8. un alphabet `stemID` pour rendre compte du nombre de radicaux allomorphiques pour une certaine catégorie flexionnelle et de leur numérotation ;
9. un alphabet `restrictchar` pour rendre compte des restrictions sur l'input des règles de réalisation autant pour des règles particulières que pour les ensembles de règles spécifiées dans des classes ou zones flexionnelles ;

---

7. Des expériences préliminaires nous ont en effet montré que l'utilisation d'alphabets distincts conduit à des descriptions plus courtes en termes de longueur de description.

10. un alphabet `translitchar` pour rendre compte des translittérations internes à Alexina<sub>PARSLI</sub> utilisées si la description comporte des symboles minimaux constitués de digraphes ;
11. un alphabet `mfschar` pour rendre compte de tout ce qui concerne les paquets de traits morphosyntaxiques, y compris les définitions des attributs et de leurs valeurs, de la définition des catégories flexionnelles et des étiquettes associées aux règles de réalisation dans la grammaire ou aux formes supplétives dans le lexique ;
12. un alphabet `operationnamechar` pour rendre compte des noms des nouvelles opérations morphologiques définies dans la description ainsi que les opérations définies par défaut dans Alexina<sub>PARSLI</sub><sup>8</sup> et effectivement employées par la description ;
13. un alphabet `operationdefchar` pour rendre compte de la définition de nouvelles opérations morphologiques.

Parmi ces alphabets, l'alphabet numéro 3 `canformchar` mesurant la longueur de description des formes canoniques n'est en général pas pris en compte dans les comparaisons entre descriptions, ces dernières ayant en effet peu de sens si elles ne sont pas faites sur le même inventaire de flexèmes. Si toutefois l'on voulait encoder les mêmes données par un choix de formes de citation différentes selon les descriptions, rien n'empêche de prendre en compte également cet alphabet dans la comparaison.

La longueur de description que nous mesurons est alors définie en généralisant la formule ci-dessus comme suit.

Si un message  $m$  est codé comme une séquence de symboles pris parmi  $p$  alphabets  $W^1, \dots, W^p$ ,  $W^1 = \{w_1^1, \dots, w_{n_1}^1\}, \dots, W^p = \{w_1^p, \dots, w_{n_p}^p\}$  de telle façon que l'alphabet dans lequel est sélectionné un symbole peut toujours être déterminé par son contexte gauche, alors nous définissons la longueur de description comme étant :

$$DL(m) = - \sum_{j=1}^p \sum_{i=1}^{n_p} o(w_i^p) \log_2 \frac{o(w_i^p)}{N_p},$$

où  $N_p$  est le nombre de symboles de l'alphabet  $W^p$  dans  $m$ .

Une telle métrique permet d'approcher la complexité d'une description structurée, mais également de mesurer la contribution de chaque alphabet à cette complexité. Cette métrique nous permet alors de calculer la longueur de description de différentes descriptions en <sub>PARSLI</sub> pour comparer ces descriptions morphologiques concurrentes pour un même (sous-)système flexionnel en termes de compacité globale. Par ailleurs, le fait que différents

8. À savoir `insert`, `replace`, `append` et `left_append`.

alphabets encodent des éléments différents de la description permet de quantifier la répartition interne de l'information entre les différents éléments de la description. En ne mesurant ainsi que le contenu des éléments encodés par les alphabets 5 et 6 (voire 10) nous pouvons ainsi isoler l'information morphologique/morphographémique relative à la flexion du reste. Par la mesure séparée de l'information contenue dans les fichiers de la grammaire d'une part et du lexique d'autre part, nous pouvons aussi par ailleurs évaluer la quantité d'information contenue dans le lexique flexionnel, et dans la grammaire flexionnelle.

Dans les trois sections suivantes nous observons plus en détail la question de la répartition de l'information relative à un système flexionnel donné selon une vision modulaire de la grammaire du type de celle que nous avons proposée en conclusion du chapitre 2 de cette thèse.

## 6.2 Répartition de l'information : lexicale ou grammaticale

Cette section s'appuie sur les expériences réalisées sur quatre descriptions concurrentes du français publiées dans (Sagot & Walther, 2011) et (Walther & Sagot, 2011)<sup>9</sup>. Nous y appliquons notre mesure de compacité développée à la section 6.1 comme critère d'évaluation de descriptions morphologiques concurrentes pour un même (sous-)système flexionnel. Cette section illustre également la question de la répartition de l'information flexionnelle entre la grammaire et le lexique flexionnels.

### 6.2.1 Représenter la flexion verbale du français

Nous avons mentionné au chapitre 4 différentes façons de décrire certains aspects de la flexion verbale du français, qui correspondent à des répartitions différentes entre les informations décrites dans le lexique et dans la grammaire morphologique. Nous en rappelons ici les deux éléments principaux qui ont joué un rôle majeur dans les expériences décrites dans cette section.

**Représentation des verbes en *-ayer*** Le premier point concerne la représentation des verbes en *-ayer*. Dans (Bonami & Boyé, 2010a, pp. 5-6), les auteurs considèrent ces verbes

---

9. Parmi ces quatre descriptions, l'une emploie explicitement des notions comme celle de *zone réalisationnelle* définies dans le modèle  $\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}$  tel que décrit dans (Walther, 2011b). Toutefois,  $\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}$  ainsi que son implémentation  $\text{Alexina}_{\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}}$  ont évolué depuis. Les quelques différences observables dans les résultats de l'expérience par rapport à  $\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}$  tel que décrit au chapitre 4 et  $\text{Alexina}_{\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}}$  telle que décrite au chapitre 5 sont la conséquences des améliorations qui leur ont été apportées depuis.

implicitement<sup>10</sup> comme polyparadigmatiques. Chacun des verbes en *-ayer* est associé à deux jeux de radicaux, l'un ne comportant qu'un radical unique en *ej*, l'autre comportant deux radicaux (en *ej* et en  $\varepsilon$ ). Chacun de ces jeux de radicaux est fléchi selon les règles de flexion des verbes du premier groupe en français. Le résultat est la génération des deux paradigmes complets pour un même verbe. Dans ces paradigmes, certaines formes (construites sur un radical en *ej*) sont par ailleurs identiques. Par exemple, la première personne du pluriel de l'indicatif *balayons* (*balejδ*) est obtenu à la fois par la combinaison du radical *balej* du premier jeu de radicaux avec le suffixe  $\delta$  et par la combinaison de *balej* du second jeu de radicaux avec  $\delta$ . Les formes *balaie* et *balaye* sont en revanche construite par l'emploi de deux radicaux différents : *bale* du premier jeu de radicaux et *balej* du second jeu, chacun associé à une absence de marquage.

Notre façon de représenter la flexion de ces verbes, employant les notions introduites dans  $\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}$ , est la suivante. Nous définissons deux radicaux, l'un en *-ay-* ou *ej* et l'autre en *-ai-* ou  $\varepsilon$ . Par ailleurs, nous partitionnons la classe flexionnelle  $Z_1^{exp}$  des verbes du premier groupe en deux zones flexionnelles : l'une, notée  $z_1^{exp}$ , utilisée seulement avec le radical en *-ay-* et l'autre, notée  $z_2^{exp}$ , utilisée par les deux radicaux. Ainsi, nous obtenons trois combinaisons de zones possibles :  $z_1^{exp}$  et le radical en *-ay-*,  $z_2^{exp}$  et le radical en *-ai-* et  $z_2^{exp}$  et le radical en *-ay-*. Dans  $\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}$ , ces trois combinaisons seront qualifiées de SOUS-SCHÈMES appartenant à un SCHÈME complet et spécifique à une classe d'entrées lexicales, en l'occurrence à celui des verbes en *-ayer*. Le rappel de notre analyse est illustré par les figures 6.2 et 6.3 pour les formes du présent de l'indicatif de BALAYER.

Classe $Z_{v1}^{exp}$		
1.SG	-e	<span style="color: blue;">■</span> $z_1^{exp}$
2.SG	-es	
3.SG	-e	<span style="color: lightblue;">■</span> $z_2^{exp}$
1PL	-ons	
2PL	-ez	
3PL	-ent	

FIGURE 6.2 : Zones flexionnelles du présent de l'indicatif des verbes du premier groupe en français

10. *La plupart des lexèmes ont un paradigme stable, chaque case du paradigme étant clairement remplie par une forme et une seule, mais il y a des exceptions. Un cas particulièrement parlant est celui du verbe asseoir en français. [...] Le cas d'ASSEOIR est particulièrement net, mais il n'est pas isolé. Le verbe HAÏR connaît les présents singuliers je hais et je haïs ; le verbe PAYER dispose de je paie et je paye ; etc. [...] Il est donc rare, mais pas impossible, qu'un lexème dispose de plusieurs paradigmes en concurrence. Cette situation constitue une exception claire à la fonctionnalité de la flexion : pour un lexème et un paquet de traits donnés, il y a plus d'une forme possible.* (Bonami & Boyé, 2010a, pp. 5-6).

	$balay + z_{v1,1}^{exp}$	$balai + z_{v1,2}^{exp}$	$balay + z_{v1,2}^{exp}$
1.SG		<i>balaie</i>	<i>balaye</i>
2.SG		<i>balaies</i>	<i>balayes</i>
3.SG		<i>balaie</i>	<i>balaye</i>
1.PL	<i>balayons</i>		
2.PL	<i>balayez</i>		
3.PL		<i>balaient</i>	<i>balayent</i>

FIGURE 6.3 : sous-schémas (simplifiés) pour le verbe *balayer*

La principale différence entre notre analyse de ces verbes et celle de Bonami & Boyé (2003) est donc que dans (Bonami & Boyé, 2003), chaque radical en *-ayer* est listé deux fois, alors que dans notre analyse il n'apparaît qu'une seule fois. Notre analyse est donc en ce point moins redondante.

**Représentation des verbes du deuxième groupe** L'autre cas essentiel est la représentation des verbes du deuxième groupe.

Si l'on utilise l'approche à douze radicaux de Bonami & Boyé (2003)<sup>11</sup>, ces verbes peuvent définir explicitement un second radical en *-iss-* (dans le lexique), à côté du radical de base en *-i-*, cf. l'exemple pour *finir* ci-dessous.

#### 6.1 Analyse avec radicaux multiples (espaces thématiques) : exemple *finir*

Thème 1 : *fni-*

Thème 2 : *fniss-*

Notre analyse fait ici usage d'une classe de radicaux spécifique. Cette solution consiste à associer les lexèmes du deuxième groupe à deux zones de radicaux appartenant à une classe de radicaux unique. Ces zones de radicaux sont la zone  $z_1^s$  pour les cases pour lesquelles les verbes du deuxième groupe utilisent un radical secondaire en *-iss-*, et la zone  $z_2^s$  pour les autres cases utilisant le radical en *-i-*. Chacune de ces zones peut alors être associée à la classe flexionnelle (unique)  $Z_2^{exp}$  des verbes du deuxième groupe. Le schéma des verbes du second groupe est alors défini au moyen de deux sous-schémas (simplifiés), à savoir la zone de radicaux  $z_1^s$  associée à la classe flexionnelle  $Z_2^{exp}$  et la zone de radicaux  $z_2^s$  combinée à  $Z_2^{exp}$ . Les formes issues de chacun de ces sous-schémas sont à chaque fois les formes correspondant aux cases définies par les paquets de traits morphosyntaxiques compatibles de chacun des éléments d'un sous-schéma donné : en d'autres termes, chaque sous-schéma définit

11. FINIR ne les distingue pas tous.

les réalisations des formes exprimant l'intersection des ensembles de paquets de traits morphosyntaxiques pour lesquels chacune de ses zones est définie. Cela évite notamment toute redondance, voire pire les cas de surgénération, dans la réalisation des formes d'un lexème donné.

Notre modélisation des verbes du deuxième groupe du français est ré-illustrée pour le présent de l'indicatif dans les figures 6.4 et 6.5.

Cl. de radicaux $Z_{v2}^s$			Cl. flexionnelle $Z_{v2}^{exp}$		
1.SG		■ $z_{v2,1}^s$	1.SG	-s	■ $Z_2^{exp}$
2.SG	-i		2.SG	-s	
3.SG			3.SG	-t	
1.PL		■ $z_{v2,2}^s$	1.PL	-ons	
2.PL	-iss		2.PL	-ez	
3.PL			3.PL	-ent	

FIGURE 6.4 : Classe et zones de radicaux (à gauche) et classe flexionnelle (à droite) du présent de l'indicatif des verbes du deuxième groupe en français

	$fini + Z_2^{exp}$	$finiss + Z_2^{exp}$
1.SG	<i>finis</i>	
2.SG	<i>finis</i>	
3.SG	<i>finit</i>	
1.PL		<i>finissons</i>
2.PL		<i>finissez</i>
3.PL		<i>finissent</i>

FIGURE 6.5 : Sous-schémas (simplifiés) pour le verbe *finir*

Pour les verbes du deuxième groupe, la différence entre les deux analyses est ainsi la suivante : chez Bonami & Boyé (2003) la différence entre le premier et le deuxième groupe est codée en listant un second radical dans le lexique pour chaque verbe alors que dans notre analyse cette différence est codée par des règles de formation de radical régulières qui s'applique à l'intégralité des verbes du deuxième groupe. Les verbes du deuxième groupe n'ont ainsi pas besoin de spécifier de second radical. On s'attend ainsi à une diminution de la longueur de description de notre analyse par rapport à celle de Bonami & Boyé.



## 6.2.2 Mesurer la complexité de différentes descriptions morphologiques de la flexion verbale du français

Pour les expériences décrites dans cette section, nous avons utilisé l'inventaire de lexèmes verbaux du lexique du français écrit *Lefff* Sagot (2010). Nous avons ignoré les distinctions entre entrées ne différant que par des propriétés non morphologiques (un même verbe peut avoir plusieurs entrées dans le *Lefff* correspondant à différents cadres de sous-catégorisation et/ou à différents sens). Nous avons ainsi obtenu un inventaire de flexèmes (ou lemmes). La version officielle du *Lefff* contient 7 820 lemmes verbaux parmi lesquels 6 966 verbes du premier groupe, 315 du deuxième groupe et 539 du troisième groupe. L'ensemble de ces verbes produisent 300 693 formes fléchies distinctes.

Dans la version de la couche morphologique d'Alexina<sup>12</sup> que nous avons utilisée dans nos expériences publiées dans (Sagot & Walther, 2011) et (Walther & Sagot, 2011), les informations morphologiques associées à une entrée lexicale étaient les suivantes :

- une forme de citation (l'infinifit pour les verbes français) ;
- un schème suivi d'une variante de schème optionnelle : si deux schèmes ne diffèrent que pour quelques cases, ils peuvent être fusionnés, des règles alternatives de réalisation sont associées à ces cases, ces règles étant déclenchées lexicalement par ces variantes de schème<sup>13</sup> ;
- facultativement, un schème radical<sup>14</sup> (ne pas spécifier le schème radical indique que l'on doit utiliser le schème radical par défaut associé au schème) ;
- facultativement, une liste de radicaux supplétifs (les radicaux non spécifiés sont à construire par l'application des règles de formation de radicaux associées au schème radical).

Ainsi, dans l'état d'Alexina<sub>PARSL</sub> de (Sagot & Walther, 2011 ; Walther & Sagot, 2011), une entrée telle que « bouillir <sub>v23r/bouill, bou</sub> » correspondait à un flexème (ou lemme) dont la forme de citation est *bouillir*, le schème *v* (avec la variante de schème<sub>23r</sub>), le schème radical par défaut pour le schème *v* tel qu'indiqué dans la grammaire morphologique, et dont le radical 1 est *bouill* et le radical 3 *bou* (les autres radicaux étant à construire à partir de ces deux radicaux explicités et des règles de formation de radicaux du schème radical).

12. Il s'agissait au moment des expériences publiées dans (Sagot & Walther, 2011) et (Walther & Sagot, 2011) du formalisme Alexina<sub>PARSL</sub> à un stade intermédiaire de son évolution. Alexina<sub>PARSL</sub> implémentait à ce moment le modèle <sub>PARSL</sub> tel que décrit dans (Walther, 2011b).

13. Ces variantes sont toujours utilisées dans les implémentations en Alexina<sub>PARSL</sub>, mais elles ont été généralisées à d'autres éléments depuis, cf. chapitre 5.

14. Le schème radical n'est plus utilisé dans la dernière version de <sub>PARSL</sub> et d'Alexina<sub>PARSL</sub>. Il correspond aux zones de radicaux désormais intégré dans le schème général.

Nous allons désormais décrire brièvement les quatre modélisations différentes de la flexion verbale du français que nous avons développées et implémentées. Pour illustrer notre propos, les entrées lexicales pour un petit ensemble de flexèmes dans chacune de ces descriptions sont données au tableau 6.1. Elles se distinguent essentiellement par leur emploi variable de l'allomorphie radicale et de classes flexionnelles multiples.

FORME DE CITATION	FLAT	ORIG	NEW	BoBo
aimer	v1	v-er <sub>std</sub>	v-er	v <sub>1</sub>
acheter	v18	v-er <sub>std</sub>	v-er	v <sub>1</sub>
jeter	v8	v-er <sub>dbl</sub>	v-eter	v <sub>1</sub> /jett <sub>,,,,,</sub> jettə
balayer	v12	v-ayer	v-ayer	v-ayer <sub>1</sub>
finir	v2	v-ir <sub>2</sub>	v-ir <sub>2</sub>	v <sub>23r</sub>
requérir	v42	v-ir <sub>3</sub>	v-ir <sub>3</sub>	v <sub>23r</sub> /requér,requier <sub>,,,,,</sub> requer,requi,requis
cueillir	v51	v-assaillir	v <sub>23r</sub> /cueill <sub>,,,,,</sub> cueillə	v <sub>23r</sub> /cueill <sub>,,,,,</sub> cueilla
prendre	v24	v-prendre	v-prendre	v <sub>3re</sub>
mettre	v17	v-mettre	v-mettre	v <sub>3re</sub> / <sub>met</sub> <sub>,,,,,</sub> mi,mis

TABLE 6.1 : Entrées lexicales pour quelques flexèmes dans chacune de nos quatre représentations concurrentes de la flexion verbale du français

**BoBo, une analyse à classe flexionnelle unique** À une extrémité du spectre se trouve l'analyse de Bonami & Boyé (2003), qui utilise seulement une classe flexionnelle mais douze radicaux. Nous sommes partis d'une implémentation préliminaire en DATR de ce modèle réalisée par Olivier Bonami<sup>15</sup>. Pour générer correctement toutes les formes surabondantes, nous avons par ailleurs dû étendre cette analyse. Le résultat est une description, appelée BoBo, qui ne contient qu'une classe flexionnelle, plusieurs schèmes (4 pour les verbes non défectifs, quelques autres pour les verbes défectifs) et 61 règles morphographémiques. Cette description repose fortement sur une caractéristique importante de notre implémentation en Alexina<sub>FRSLI</sub> : toute information non explicitée est reconstruite par des mécanismes par défaut. Par exemple, ne pas spécifier un certain radical dans une entrée lexicale conduit à utiliser son schème radical pour le produire<sup>16</sup> ; par ailleurs, dans la version d'Alexina<sub>FRSLI</sub> utilisée dans (Sagot & Walther, 2011 ; Walther & Sagot, 2011), si le schème radical n'est pas spécifié, on utilise le schème radical par défaut associé au schème de l'entrée ; si ce schème ne définit pas de schème radical par défaut, on utilise le même radical que celui de la forme de citation. Le résultat est que BoBo a une longueur de description d'environ

15. Parce que cette analyse était implémentée sur les phonèmes et non les sur graphèmes, nous avons dû appliquer un certain nombre de transformations pour nous permettre de représenter sans erreur la flexion graphémique, y compris par l'ajout de règles de sandhi.

16. Autrement dit cela revient à utiliser les règles de construction de radical régulières.

51 400 bits qui se répartissent en 46 000 bits dans le lexique (ce qui est particulièrement élevé et provient du fait que de nombreux radicaux sont spécifiés explicitement<sup>17</sup> et 5 400 bits dans la grammaire morphologique, ce qui est très bas, comme attendu.

**FLAT, une analyse plate sans allomorphie radicale** À l'autre extrémité du spectre de descriptions possibles est une analyse que nous avons générée automatiquement. Nous avons produit une description « plate », nommée FLAT, qui n'utilise pas les notions d'allomorphie radicale, de règles morphographémiques et de zones flexionnelles, de la façon suivante. Nous avons extrait pour chaque verbe la sous-chaîne commune la plus longue entre toutes ses formes et l'avons considérée comme son radical. Dans chacune de ces formes, ce qui est à droite de cette partie commune a été considéré comme la « désinence » de la forme. Chaque verbe a alors reçu une *signature* qui est la liste des désinences ainsi extraites, ordonnées par ordre alphabétique de l'étiquette morphosyntaxique correspondante. Ensuite, on attribue à tous les verbes ayant une même signature une même classe flexionnelle définie à partir de cette signature. Le résultat est une description comportant 139 classes flexionnelles. Sa longueur de description, mesurée de la façon expliquée à la section précédente, est d'environ 131 400 bits (9 200 bits dans le lexique<sup>18</sup> et 122 200 bits dans la grammaire morphologique).

**ORIG, la description originale du Lefff** Entre ces deux extrêmes, la description ORIG utilisée actuellement par le Lefff, qui repose sur un nombre très important de règles morphographémiques mais utilise 92 classes flexionnelles dont la majorité pour les verbes du troisième groupe, a une longueur de description de 82 900 bits (8 100 bits dans le lexique, 74 800 bits dans la grammaire).

**NEW, une analyse en PARSLI** Mais c'est une autre description qui donne les résultats les plus intéressants. Comme indiqué plus haut, en utilisant la notion de zone flexionnelle conjointement à un ensemble raisonnable de règles morphographémiques, nous avons pu développer une description morphologique plus satisfaisante des verbes français, nommée NEW, qui utilise 20 schèmes. Parmi eux, se trouvent notamment un schème pour les verbes

---

17. Comme le montre l'exemple de FINIR dans le tableau 6.1, notre implémentation de l'analyse de Bonami & Boyé (2003) a en fait même compacté l'approche d'origine des auteurs en ne spécifiant pas dans le lexique les radicaux pour les nombreux verbes du deuxième groupe, mais en les générant par la grammaire. On s'attendrait à ce qu'une implémentation plus fidèle augmente encore davantage la longueur de description globale de BoBo.

18. Ici et dans tous les résultats qui suivent, la contribution de la liste de formes de citation est ignorée, puisqu'elle est commune à toutes les descriptions.

du premier groupe sans surabondance, un schème pour les verbes en *-ayer* en utilisant les zones  $z_1^{exp}$  et  $z_2^{exp}$  décrites à la section 6.2.1 et un schème pour les verbes du deuxième groupe, analysés comme réguliers au moyen de deux sous-schémas, comme indiqué à cette même section. Les schèmes définis pour les verbes du troisième groupe reprennent les classes flexionnelles les plus peuplées de ORIG (verbes en *-ir*, verbes en *-re* les « moins irréguliers », verbes en *-prendre*, *-mettre*, etc), et traitent les autres verbes comme dans BoBo. La longueur de description correspondante, seulement 35 800 bits, est inférieure à celle de BoBo. Elle se répartit en 20 100 bits dans le lexique (deux fois plus que pour FLAT, mais deux fois moins que pour BoBo) et 15 700 bits dans la grammaire morphologique (trois fois plus que dans BoBo, mais huit fois moins que dans FLAT).

**Résultats** Les caractéristiques principales des quatre descriptions sont résumées au tableau 6.6. Les différents résultats des mesures de Longueur de description Minimale pour chacune des quatre descriptions, indiqués en ordonnant les descriptions le long du spectre, sont représentés à la figure 6.7. Ils montrent bien que l'utilisation de la notion de zone flexionnelle introduite dans  $\mathcal{PARSLI}$ , permet de rendre compte de la morphologie verbale du français d'une façon plus économique en termes de longueur de description et donc de compacité par rapport aux descriptions traditionnelles mais aussi par rapport à une description récente et originale Bonami & Boyé (2003). On peut noter que nous n'aurions pas obtenu ce résultat si nous n'avions pas pris en compte la longueur de description du lexique, mais seulement celle de la grammaire morphologique. Cela dit, puisque la répartition de l'information morphologique entre lexique et règles varie d'une description à une autre, cela n'aurait pas grand sens d'évaluer la longueur de description de la grammaire morphologique seule.

	RADICAUX	RÈGLES MORPHOGRA- PHÉMIQUES	NOTIONS DE $\mathcal{PARSLI}$	NOTIONS DE FACTORISATION	CLASSES. FLEXION- NELLES	ENTRÉES LEXICALES
ORIG	1	oui	oui	non	92	simples
FLAT	1	non	non	non	139	simples
BoBo	12	non	oui	partiel	1	complexes
NEW	1-12	oui	oui	oui	20	moyennes

FIGURE 6.6 : Caractéristiques principales des quatre descriptions de la morphologie verbale du français

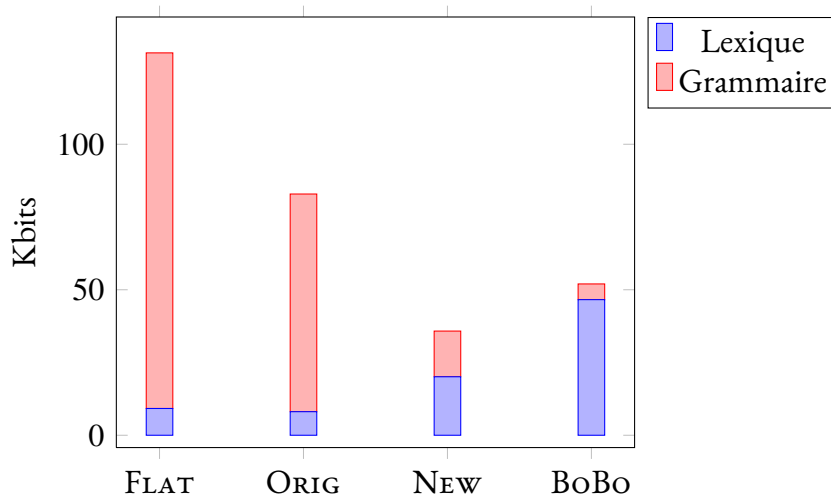


FIGURE 6.7 : Longueurs de description de différentes descriptions de la morphologie verbale du français

### 6.2.3 Compacité et formalisation

Dans les expériences proposées dans (Sagot & Walther, 2011) et (Walther & Sagot, 2011) et décrites dans cette section, nous avons étudié la compacité de quatre descriptions concurrentes de la flexion verbale du français<sup>19</sup> : la description de (Bonami & Boyé, 2003), une description plate générée automatiquement, la description originale du *Lefff* officiel et une description nouvelle faisant usage des notions descriptives introduites dans  $\mathcal{PARSLI}$ . Les mesures de compacité que nous avons obtenues ont été calculées à l'aide de la métrique introduite à la section 6.1 issue de la théorie de l'information et adaptée à notre modèle et notamment au caractère structuré des descriptions. Ces mesures nous ont apporté deux résultats théoriques principaux.

D'une part, les résultats quantitatifs obtenus sur nos quatre descriptions montrent que des façons traditionnelles de décrire la morphologie verbale du français au moyen de nombreuses classes flexionnelles, mais également un modèle récent et radicalement différent Bonami & Boyé (2003), sont tous d'une complexité plus élevée qu'une description que nous avons développée en utilisant des notions que nous avons introduites dans  $\mathcal{PARSLI}$ . En particulier les notions de zone flexionnelle et de schème permettent d'équilibrer au mieux la répartition des informations flexionnelles entre le lexique et la grammaire flexionnelles.

Cependant, ce gain d'information n'est visible que dès lors que l'on adopte une vision

19. Les quatre descriptions dans le format intermédiaire d'Alexina $\mathcal{PARSLI}$  sont disponibles en ligne sur la page personnelle de l'auteur.

modulaire intégrée de la grammaire telle que celle que nous avons présentée en conclusion du chapitre 2. Déplacer une quantité d'information d'un module de la grammaire vers un autre augmentera nécessairement la compacité du premier et diminuera celle du second. La compacité globale de la description dépendra de la répartition optimale entre les différents modules de la grammaire. Ainsi, comme le montre la figure 6.7, la description FLAT comporte-t-elle la quasi-totalité de l'information flexionnelle dans la composante de la grammaire flexionnelle, tandis que la description BOBO a une grammaire très compacte mais un lexique nettement moins compact. Les deux descriptions génèrent les mêmes formes, mais l'un dans l'autre, BOBO est significativement plus compacte que FLAT et, selon le principe du rasoir d'Okham, préférable à cet égard. Toutefois, il est possible d'améliorer encore la compacité de la description globale de la flexion verbale française. Notre description NEW faisant usage des nouvelles notions introduites dans PARSLI permet d'optimiser encore plus la répartition de l'information entre lexique flexionnel et grammaire flexionnelle, menant ainsi à la description la plus compacte parmi celles sur lesquelles nous avons fait nos expériences.

Notons enfin que les mesures que nous avons proposées sont des mesures du taux de redondance brut dans la distribution entre lexique et grammaire flexionnelles. Ces mesures ne se veulent pas une modélisation de la manière dont l'information est effectivement stockée par un locuteur réel du français. En particulier, elles ne disent rien ni du fonctionnement de la mémoire d'un locuteur ni de l'efficacité pour ce dernier d'un stockage particulier pour l'accès rapide à l'information. Ce type de mesures nécessiterait une batterie d'expériences d'une autre nature qui ne fait pas l'objet de cette thèse.

### 6.3 Filtrer l'information : morphologie vs. phonologie

À la section précédente, nous avons étudié la question de la distribution optimale de l'information flexionnelle entre la grammaire flexionnelle et le lexique. Toutefois, le lexique flexionnel n'est pas la seule composante de la langue à être en interaction avec la morphologie flexionnelle. Si nous adoptons une vision de la langue du type de celle que nous avons esquissée en conclusion du chapitre 2, nous nous devons d'étudier, en plus de son interface avec le lexique, au moins les interfaces de la grammaire flexionnelle avec la morphologie et la syntaxe.

Dans cette section, nous nous intéressons ainsi d'abord à la délimitation entre morphologie et morphonologie ainsi qu'à la distribution de l'information grammaticale entre ces deux composantes. En prenant l'exemple d'une analyse de la flexion du maltais

concurrente de celle que nous avons proposée au chapitre 5, nous montrons que l'analyse purement morphologique que nous avons présentée à la section 5.3 du chapitre 5 peut être en partie prise en charge par la composante morphologique<sup>20</sup>.

### 6.3.1 Analyse phonologique des radicaux maltais

#### 6.3.1.1 Réduire la complexité morphologique

Au chapitre 5 et à la suite de (Camilleri, 2011, p. 1 et p. 83), nous avons présenté une analyse des verbes maltais qui conçoit les rapports entre les radicaux d'un paradigme donné comme le résultat d'une fonction morphologique au sens d'Aronoff (1994).

Dans cette section, nous montrons que les propriétés des classes de syncrétisme radical (CSR) posées par Camilleri pour rendre compte de l'allomorphie radicale en maltais ne sont pourtant pas des constructions purement morphologiques, puisqu'une partie importante de la complexité du système peut aussi bien être résolue par la composante phonologique ou morphologique de la langue.

Nous montrons que la réalisation des formes verbales n'est pas la concaténation d'un ensemble d'affixes et de radicaux dont l'allomorphie dérive d'une morphologie autonome de haute complexité ; au contraire, la majeure partie de la complexité des CSR se réduit aisément à un simple jeu de contraintes phonologiques et morphologiques<sup>21</sup>. L'allomorphie radicale et la structure prosodique des radicaux<sup>22</sup> peuvent être déduites de règles phonologiques et morphologiques déclenchées par l'affixation. En d'autres termes, la forme de l'affixe et sa position (préfixe ou suffixe) influent sur la structure générale du mot, induisant ainsi des changements phonologiques qui modifient de façon prédictible la forme des radicaux et sont à l'origine des CSR observées par Camilleri (2011). Seuls les changements vocaliques sont effectivement imprédictibles, et doivent donc résulter de véritables règles morphologiques<sup>23</sup>.

Nous proposons donc une liste de douze règles et contraintes et stratégies de réparation

---

20. Même dans cette nouvelle analyse, il reste toutefois des données qui ne peuvent être traitées qu'en morphologie et qui constituent donc de la morphologie autonome au sens d'Aronoff (1994).

21. L'approche que nous suivons ici pour rendre compte des mécanismes morphologiques s'inspire librement d'une approche de type *contrainte et stratégies de réparation* telle que développée par Paradis (1988). Ce choix ne correspond en aucun cas d'un parti pris théorique fort. D'autres approches comme la théorie de l'Optimalité (*OT* (Prince & Smolensky, 1993)) ou comme celle de Plénat (1987) auraient également pu être employées ici.

22. En termes de structures syllabiques et de placement de l'accent.

23. Dans l'analyse complète de la flexion du premier binyan maltais que nous donnons au chapitre 7, nous reviendrons plus explicitement sur les règles régissant les changements vocaliques.

phonologiques<sup>24</sup> (et morphologiques) qui permettent de prédire, aux changements vocaliques près, l'ensemble des CSR observables et de construire les formes des paradigmes verbaux maltais.

Notons que les contraintes phonologiques se distinguent des contraintes morphologiques en ce qu'elles s'appliquent nécessairement à l'intégralité des séquences phoniques de la langue et ne sont par définition pas restrictibles à une (sous-) classe de lexèmes de la langue. Les règles morphologiques se distinguent des règles phonologiques en ce que, contrairement à ces dernières, elles sont sensibles à la structure morphotactique. Celles que nous utiliserons ici partagent néanmoins avec les règles phonologiques au sens strict le fait de s'appliquer à l'intégralité du système de la langue et de ne pas admettre de restrictions par rapport à une classe ou sous-classe de lexèmes.

Les règles et contraintes (mor)phonologiques que nous identifions ici permettent de réduire très largement la complexité indiquée en morphologie : au lieu du système de CSR développé par Camilleri (2011), nous verrons qu'il suffit de stipuler, selon les classes de lexèmes, au plus certains changements vocaliques pour rendre compte des alternances de radicaux en maltais. Le reste de la complexité est résolue par un ensemble de douze contraintes phonologiques et morphologiques décrit ci-dessous. Nous donnons à la section 6.3.2 une évaluation quantitative de la compacité respective de ces deux descriptions à l'aide de la métrique que nous avons introduite à la section 6.1.

Nous rappelons au tableau 6.2 cinq paradigmes de verbes appartenant à des bases verbales différentes selon la classification de Camilleri (2011). La classe flexionnelle unique que nous avons proposée pour rendre compte de la flexion du maltais est indiquée au tableau 5.7 du chapitre 5, page 315.

Nous rappelons également au tableau 6.3 le placement de l'accent sur les formes verbales des cinq bases verbales définies par Camilleri (2011).

### 6.3.1.2 Règles phonologiques

Nous présentons en un premier temps un inventaire de douze contraintes et stratégies de réparation phonologiques et morphologiques appartenant au système phonologique du maltais. Elles se répartissent en trois types : celles définissant les suites ou types de suites de phonèmes licites, celles sur les structures syllabiques et celles concernant le placement

---

24. La phonologie n'étant pas le sujet central de cette thèse, nous n'entrons ici volontairement pas dans le débat sur la différence entre RÈGLES et CONTRAINTES. Dans la suite de cette section, nous considérons comme RÈGLES (MOR)PHONOLOGIQUES tout ce qui provoque des changements (mor)phonologiques sur les formes, et comme CONTRAINTES (MOR)PHONOLOGIQUES ce qui provoque ou bloque l'application desdites règles.



	KITEB « écrire »	BENA « construire »	SAR « devenir »	LAQAT « frapper »	HATAF « saisir »	MESS « toucher »
PFV 1.SG	<i>ktib-t</i>	<i>bnej-t</i>	<i>sir-t</i>	<i>lqat-t</i>	<i>ħtaf-t</i>	<i>messej-t</i>
PFV 2SG	<i>ktib-t</i>	<i>bnej-t</i>	<i>sir-t</i>	<i>lqat-t</i>	<i>ħtaf-t</i>	<i>messej-t</i>
PFV 3SG.M	<i>kiteb</i>	<i>bena</i>	<i>sa:r</i>	<i>laqat</i>	<i>ħataf</i>	<i>mess</i>
PFV 3SG.F	<i>kitb-et</i>	<i>bnie-t</i>	<i>sa:r-et</i>	<i>laqt-et</i>	<i>ħatf-et</i>	<i>mess-et</i>
PFV 1PL	<i>ktib-na</i>	<i>bnej-na</i>	<i>sir-na</i>	<i>lqat-na</i>	<i>ħtaf-na</i>	<i>messej-na</i>
PFV 2PL	<i>ktib-t-u</i>	<i>bnej-t-u</i>	<i>sir-t-u</i>	<i>lqat-t-u</i>	<i>ħtaf-t-u</i>	<i>messej-tu</i>
PFV 3PL	<i>kitb-u</i>	<i>bne:-w</i>	<i>sa:r-u</i>	<i>laqt-u</i>	<i>ħatf-u</i>	<i>messe:-w</i>
IPFV 1.SG	<i>n-i-kteb</i>	<i>n-i-bni</i>	<i>n-si:r</i>	<i>n-o-lqot</i>	<i>n-a-ħtaf</i>	<i>n-miss</i>
IPFV 2SG	<i>t-i-kteb</i>	<i>t-i-bni</i>	<i>s-si:r</i>	<i>t-o-lqot</i>	<i>t-a-ħtaf</i>	<i>t-miss</i>
IPFV 3SG.M	<i>j-i-kteb</i>	<i>j-i-bni</i>	<i>j-si:r</i>	<i>j-o-lqot</i>	<i>j-a-ħtaf</i>	<i>j-miss</i>
IPFV 3SG.F	<i>t-i-kteb</i>	<i>t-i-bni</i>	<i>s-si:r</i>	<i>t-o-lqot</i>	<i>t-a-ħtaf</i>	<i>t-miss</i>
IPFV 1PL	<i>n-i-ktb-u</i>	<i>n-i-bn-u</i>	<i>n-si:r-u</i>	<i>n-o-lqt-u</i>	<i>n-a-ħtf-u</i>	<i>n-miss-u</i>
IPFV 2PL	<i>t-i-ktb-u</i>	<i>t-i-bn-u</i>	<i>s-si:r-u</i>	<i>t-o-lqt-u</i>	<i>t-a-ħtf-u</i>	<i>t-miss-u</i>
IPFV 3PL	<i>j-i-ktb-u</i>	<i>j-i-bn-u</i>	<i>j-si:r-u</i>	<i>j-o-lqt-u</i>	<i>j-a-ħtf-u</i>	<i>j-miss-u</i>
IMP 2SG	<i>i-kteb</i>	<i>i-bni</i>	<i>si:r</i>	<i>o-lqot</i>	<i>a-ħtaf</i>	<i>miss</i>
IMP 2PL	<i>i-ktb-u</i>	<i>i-bn-u</i>	<i>si:r-u</i>	<i>o-lqt-u</i>	<i>a-ħtf-u</i>	<i>miss-u</i>

TABLE 6.2 : Paradigmes de verbes du premier binyan

de l'accent. Elles sont également résumées sous forme de liste numérotée à la fin de cette section. Par ailleurs, ces contraintes sont toutes plus ou moins fortes, de sorte que certaines peuvent annuler l'influence d'autres, tandis que les stratégies de réparation sont soumises au jeu des contraintes. Nous définissons une hiérarchie (externe) des contraintes à la figure 6.10.

**6.3.1.2.1 Contraintes sur les suites de phonèmes** Le maltais écrit comporte un inventaire graphémique de 24 consonnes et de 5 voyelles. Ces graphèmes représentent 25 consonnes et 11 voyelles phonémiques (cinq longues, cinq courtes et une diphtongue). L'inventaire est donné dans le tableau de correspondance graphème-phonème 6.4 et dans le récapitulatif des contraintes phonémiques ci-dessous. La correspondance entre graphèmes et phonèmes est quasiment bijective<sup>25</sup>. En termes d'analyse, l'emploi d'une convention de notation par rapport à l'autre peut ainsi être considéré comme équivalent. Par convention, nous utilisons donc l'inventaire graphémique dans la suite de cette section.

25. Certaines consonnes comme le <x> et le <z> peuvent présenter des variations de prononciation par assimilation liées à leur placement ; une consonne <gh>, provenant du substrat sémitique de la langue, a perdu sa réalisation phonémique et est en maltais réduite à un petit coup de glotte ou un allongement de la voyelle adjacente, voire pas prononcée du tout. De même, le le <h> est parfois prononcé comme un simple « h » aspiré, parfois non prononcé.

	CVVCVC	CVCCVC	CVCV	CVCC	CV:C	
PFV	1.SG	CCV́Ct	CCV́jt	CVC.Céjt	CV́Ct	
	2SG	CCV́Ct	CCV́jt	CVC.Céjt	CV́Ct	
	1PL	CCV́C.na	CCV́j.na	CVC.Céj.na	CV́C.na	
	2PL	CCV́C.tu	CCV́j.tu	CVCCéj.tu	CV́C.tu	
	3.M.SG	CV́.CVC	CV́.CV	CV́CC	CV́:C	
	3.F.SG	CV́C.Cet	CCV́:t/CCít	CV́C.Cet	CV́:.Cet	
	3PL	CV́C.Cu	CCV́w	CVCCéw	CV́:.Cu	
	IPFV	1.SG	nV́C.CVC	nV́C.CV	nCV́CC	nCV́:C
		2SG	tV́C.CVC	tV́C.CV	tcV́CC	tcV́:C
3.M.SG		jV́C.CVC	jV́C.CV	jcV́CC	jcV́:c	
3.F.SG		tV́C.CVC	nV́C.CV	tcV́CC	tcV́:C	
1PL		nV́C.CCu/nV́.CV́C.Cu	nV́C.Cu/nVC.CV́w	nCV́C.Cu	nCV́:.Cu	
2PL		tV́C.CCu/tV́.CV́C.Cu	tV́C.Cu/tVC.CV́w	tcV́C.Cu	tcV́:.Cu	
3PL		jV́C.CCu/jV́.CV́C.Cu	jV́C.Cu/jVC.CV́w	jcV́C.Cu	jcV́:.Cu	
IMP		SG	V́C.CVC	V́C.CV	CV́CC	CV́:C
		PL	V́C.CCu/V́.CV́C.Cu	V́C.CCV́w/V́C.Cu	CV́CC	CV́:C

TABLE 6.3 : Structures des radicaux dans les paradigmes du premier binyan

Il est toutefois à noter que les voyelles longues ne sont pas explicitement marquées en maltais écrit. La longueur vocalique étant essentielle pour l'analyse, nous les marquons par le symbole API « : » correspondant. En dehors du marquage de la longueur des voyelles, nous suivons cependant les conventions graphémiques du maltais habituellement employées dans la plupart des études linguistiques existantes de la langue. Notons que parmi les consonnes phonémiques du maltais deux sont graphémiquement représentées par des digraphes (*gh* et *st*), de même que la diphtongue *ie*.

Le tableau 6.4 résume les graphèmes du maltais (colonnes de gauche) avec leurs phonèmes correspondants (colonnes de droite).

b	b	n	n	a	a
ċ	ʧ	p	p	e	ɛ
d	d	q	ʔ	i	ɪ
f	f	r	r	o	ɔ
ġ	ʤ	s	s	u	ʊ
g	g	t	t		
gh	( <sup>h</sup> /:)	st	st̄	a	a:
ħ	ħ	v	v	e	ɛ:
h	-(ħ)	w	w	i	ɪ:
j	j	x	ʃ/ʒ	o	ɔ:
k	k	ż	z	u	ʊ:
l	l	z	ts̄/dz̄		
m	m			ie	i <sup>ɛ</sup>

TABLE 6.4 : Correspondance entre graphèmes et phonèmes en maltais

Parmi les consonnes, deux sous-ensembles ont une influence phonotactique sur les phonèmes avoisinants : les consonnes arrières (*gh*, *ħ*, *h*, *k*, *q*) d'une part et les consonnes sonantes (*gh*, *l*, *m*, *n*, *r*) de l'autre. Nous notons respectivement par les symboles **B** et **R** les consonnes arrières et sonantes. Nous notons par ailleurs  $\bar{B}$  et  $\bar{R}$  respectivement les consonnes appartenant aux ensembles complémentaires de ceux auxquels appartiennent respectivement les consonnes arrières **B** et les consonnes sonantes **R**.

Comme indiqué plus haut, les voyelles peuvent quant à elles être courtes (*a*, *e*, *i*, *o*, *u*) ou longues. Les voyelles longues sont soit intrinsèquement longues (*a:*, *e:*, *i:*, *o:*, *u:*), soit longues par leur nature de diphtongue (*ie*). Les voyelles suivies d'une semi-voyelle comme *j* ou *w* comptent, comme les diphtongues, également comme des voyelles longues.

Les voyelles sont soumises à la contrainte suivante. Deux voyelles ne peuvent se suivre

directement. Lorsque deux voyelles se retrouvent côte à côte, une stratégie de réparation fait que la seconde provoque, selon les cas, le simple allongement de la première et tombe, ou bien fusionne avec elle en la diphtongue correspondante.

En ce qui concerne les consonnes, notons d'abord que le maltais standard tend à éviter l'apparition d'une séquence de deux consonnes CC où la seconde est une consonne sonante R sans que la première ne le soit. Il n'y a pas pour autant d'impossibilité phonétique systématique à l'existence d'une telle séquence RR. Le fait que cette séquence soit cependant problématique a provoqué l'existence de formes différentes à la troisième personne pour les verbes de base verbale CVCVC (selon la classification de Camilleri (2011)) ayant un radical dont la deuxième consonne est une sonante de verbes. Mais l'existence de telles formes alternatives au PL.IPFV des verbes à base CVRVC n'est pas partagée par tous les dialectes du maltais. Camilleri & Vanhove (1994) montrent notamment que le dialecte mgarr parlé au nord-ouest de l'île de Malte ne distingue pas entre les formes avec consonne sonante et les autres. Le fait que cette distinction soit absente d'une partie des dialectes donne ainsi à penser que plus qu'une interdiction phonotactique, elle reflète une idiosyncrasie de certains systèmes morphologiques dialectaux<sup>26</sup> et doit ainsi être stipulée à part dans la composante morphologique. Toutefois, nous pouvons noter ici que c'est cette préférence phonologique du MS, ainsi que de certains autres dialectes, qui est à l'origine d'une flexion particulière d'un sous-ensemble de verbes au pluriel de l'imperfectif.

Les contraintes phonémiques dépendent pour certaines de la position des phonèmes dans la syllabe ou le mot. Ainsi, le maltais ne possède-t-il pas de séquence de trois consonnes.

Les contraintes peuvent par ailleurs imposer l'application des RÈGLES MORPHONOLOGIQUES en guise de stratégies de réparation. Ces dernières sont déclenchées par la combinaison de deux morphes ayant chacun une structure phonologique particulière.

1. Si un morphe se termine par une voyelle et que le prochain commence par une consonne, alors la voyelle se palatalise, donnant lieu à une séquence vjC. La séquence vj prend alors les propriétés d'une diphtongue.
2. Si un morphe se termine par une voyelle et que le prochain commence par une voyelle, les règles suivantes s'appliquent :
  - a) Si la première voyelle est un *i* et la seconde un *e* : dans ce cas, *i* et *e* fusionnent en la diphtongue *ie*.
  - b) Si la seconde voyelle est un *u*, le *u* devient *w* pour donner la lieu à une séquence *vw*.

---

26. Dont le MS.

c) Dans tous les autres cas, la seconde voyelle tombe et provoque l'allongement de la première voyelle<sup>27</sup>.

3. Lorsqu'un morphe se termine par deux consonnes et que le suivant commence par (au moins) une consonne, la séquence *ej* est insérée à la frontière de morphes<sup>28</sup>.

**6.3.1.2.2 Contraintes sur les structures syllabiques** Les mots du maltais ont en général deux syllabes dont une, en général la première, est toujours accentuée.

Considérons un modèle de la SYLLABE classique dans lequel toute syllabe  $\sigma$  est constituée d'une ATTAQUE A et d'une RIME R, elle-même constituée d'un NOYAU (NUCLEUS) N et d'une CODA C. En maltais seul le noyau est obligatoire ; le noyau est toujours vocalique. La présence d'une attaque (au moins simple<sup>29</sup>) est une tendance forte. Dans les formes verbales que nous étudions, le seul cas de syllabe à initiale vocalique observable concerne les formes impératives des verbes comportant une voyelle d'extension à l'imperfectif.

Les syllabes admises en maltais ne recouvrent pas l'intégralité des combinaisons possibles. . Seules sont ainsi admises les structures syllabiques suivantes : CV, CVC, CCV, CCVC, CVCC, CCVCC, CV:C, CCV:, CCV:C (et VC<sup>30</sup>).

Nous notons  $\sigma_+$  les syllabes lourdes : CVCC, CCVCC, CV:C, CCV:, CCV:C.

Plus la coda est longue, plus la syllabe est lourde. Une voyelle longue alourdit également la syllabe. En termes de lourdeur, la longueur de la voyelle prime sur la longueur de la coda. Ainsi, les syllabes se classent de la plus légère à la plus lourde comme indiqué par la figure 6.8.

Les autres types de syllabes sont des syllabes légères. Nous les notons  $\sigma_-$  : CV, CVC, CCV, CCVC (et VC).

Les mots sont soumis aux contraintes et règles syllabiques suivantes.

1. Un mot comporte deux syllabes sauf si des contraintes plus fortes imposent la réduction du nombre de syllabes à une syllabe unique ou à la présence d'une syllabe supplémentaire (cf. plus bas).
2. En général, si l'affixation provoque l'apparition de syllabes surnuméraires, une

---

27. Cette règle par défaut implémente un cas particulier de la ELSEWHERE CONDITION formulée par Kiparsky (1973).

28. Compte tenu des désinences verbales et de la structure des verbes du premier binyan, le cas d'une frontière de morphe entre une consonne à gauche et deux consonnes à droite ne se pose pas.

29. C'est-à-dire constituée d'une seule consonne. La seconde consonne est facultative. Une attaque comportant deux consonnes est dite *double*.

30. Rare.

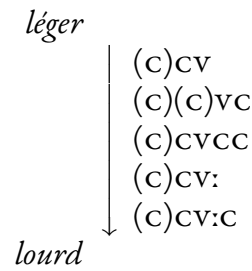


FIGURE 6.8 : Poids des syllabes maltaises

stratégie de réparation est déclenchée : des effacements de voyelles rétablissent <sup>31</sup> le nombre requis de deux syllabes par mot <sup>32</sup>. Ces effacements se font en priorités sur les voyelles à gauche de la syllabe accentuée en commençant par la voyelle la plus proche, cf. l'illustration à la figure 6.9 : l'effacement se fait dans l'ordre -1, puis -2 à n, puis 1 et enfin 2 à n, en fonction de ce qui est possible.

3. Cependant, tous les effacements de voyelles sont empêchés s'ils devaient entraîner l'apparition de séquences phonémiques interdites, à savoir les suites \*CCC. Dans ce cas, un mot peut être contraint à avoir plus de deux syllabes.

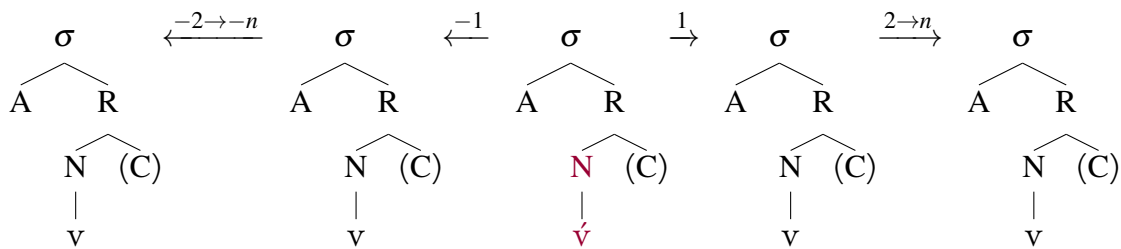


FIGURE 6.9 : Ordre préférentiel d'effacement des voyelles surnuméraires dans l'ordre suivant : -1, -2 → -n, 1, 2 → n

### 6.3.1.2.3 Contraintes d'accentuation

1. Il y a une et une seule syllabe accentuée par mot.
2. Une syllabe finale ne peut être accentuée que si elle est également l'unique syllabe du mot en question. Stratégie de réparation en cas de violation : si d'autres syllabes la précèdent, leur voyelle est effacée tant que l'effacement ne contrevient pas aux

31. Ces mécanismes fonctionnent donc comme des stratégies de réparation similaires à celles proposées par Paradis (1988).

32. Les règles sur les suites phonémiques licites et l'accentuation peuvent empêcher la suppression de certaines voyelles.

contraintes sur les suites de phonèmes énoncées plus haut (cf. les règles 1a à 1b ci-dessous).

3. Chaque lexème comporte un accent lexical qui peut éventuellement être déplacé en réaction à certains cas d'affixation. Ainsi, le lexème  $\text{HÁTA}\text{F}$  « saisir » comporte un accent lexical sur sa première position vocalique *a* (entre *h* et *t*). En revanche, dans la forme de la première personne du singulier du perfectif *h̄táft̄*, l'accent se retrouve sur la deuxième position vocalique *a* (entre *t* et *f*), tandis qu'à l'imperfectif on le retrouve sur la voyelle thématique, cf. la première personne du singulier *náhtaf* (avant *h̄*).
4. Quand un préfixe ou certains des suffixes sont ainsi combinés avec un radical donné, ils ont pour conséquence de provoquer le déplacement de l'accent du mot d'une syllabe en leur direction, c'est à dire d'une syllabe vers la gauche pour les préfixes et d'une syllabe vers la droite pour les suffixes. Ce déplacement de l'accent peut ensuite se répercuter sur des effacements de voyelles visant à satisfaire les contraintes d'accentuation ci-dessus.

En revanche, certains suffixes ne provoquent pas de déplacement de l'accent. Il s'agit des suffixes commençant par une voyelle. Ainsi, les suffixes du pluriel en *-u* et le suffixe de la troisième personne du féminin singulier du perfectif *-et* ne provoquent pas le déplacement de l'accent vers la droite. Le suffixe du pluriel étant le seul susceptible d'apparaître avec un des préfixes<sup>33</sup>, il ne peut y avoir de conflit entre le déplacement vers la gauche (préfixal) et vers la droite (suffixal pour les suffixes commençant par une consonne).

5. L'accent ne se déplace pas non plus si la syllabe sur laquelle il se trouve est plus lourde que celle sur laquelle il se retrouverait suite au déplacement.
6. Une voyelle longue est toujours accentuée.

**6.3.1.2.4 Remarque sur les changements vocaliques** Parmi les changements vocaliques pouvant affecter les radicaux du maltais, il en est un qui est clairement contraint phonologiquement. À chaque fois qu'un radical est combiné avec un des suffixes pouvant entraîner un déplacement de l'accent vers la droite, toute voyelle *e* de seconde position dans la forme de citation est changée en *i*. Les autres changements vocaliques semblent plus arbitraires phonologiquement. Nous n'avons d'autre choix à ce stade que de les traiter en morphologie.

Toutefois, l'observation sur la régularité du changement vocalique *e* vers *i* pour la seconde position vocalique reste une caractéristique marginale au sein de l'ensemble des

---

33. Les préfixes sont tous des marques de l'imperfectif.

changements vocaliques observables que nous avons listés dans les tableaux 5.13 à 5.15 au chapitre 5. Dans l'analyse morphologique que nous proposons dans cette section, nous n'avons par conséquent pas intégré cette régularité dans le système de règles et de contraintes morphologiques mis en place pour rendre compte de l'allomorphie radicale verbal du maltais. Son intégration reste cependant une option descriptive possible.

### 6.3.1.3 Hiérarchies des contraintes et des stratégies de réparation

Comme indiqué au fur et à mesure dans ce qui précède, les conflits entre contraintes sont arbitrés par un ordre extrinsèque. Nous résumons et numérotons ci dessous les différentes contraintes énoncées jusqu'ici.

Les contraintes sur les suites de phonèmes *1a* et *1b* sont absolues. Les autres contraintes ne sont respectées que dans la mesure où elle ne contreviennent pas à *1a* à *1b*. La contrainte *1c* ne constitue qu'une tendance générale, mais elle n'a pas le statut de règle phonologique en tant que telle.

**Contraintes phonologiques sur les suite de phonèmes en maltais :**

- 1a* Deux voyelles différentes ne peuvent se suivre sans fusionner en la diphtongue correspondante ou provoquer l'allongement de la première.
- 1b* Il ne peut y avoir trois consonnes de suite au sein d'une même syllabe : \*CCC.
- 1c* Le maltais standard tend à éviter les séquences de deux consonnes RR où la seconde est une consonne sonante R alors que la première ne l'est pas.

L'application des règles morphologiques intervenant aux frontières de morphes est également systématique. La règle *2b* a en particulier pour fonction de garantir le respect de la contrainte *1b* ci-dessus. De même, les règles *2e* et *2d* permettent de garantir la



contrainte 1a.

Règles morphologiques intervenant aux frontières de morphes :

2a  $\emptyset \rightarrow j / v\#\_c$

2b  $\emptyset \rightarrow ej / cc\#\_c$

Voyelles :

2c  $i\#e \rightarrow ie$

2d  $v\#u \rightarrow vw$

2e  $v_1 \rightarrow v_1: / \_ \#v_{\neq u}$  (ailleurs)

Les autres contraintes dépendent des règles et contraintes précédentes. Leur application ne peut se faire que dans la limite du respect des contraintes sur les phonèmes.

Contraintes sur la structure syllabique en maltais :

3a Chaque mot comporte deux syllabes (sauf contrainte plus forte).

Les voyelles des syllabes surnuméraires sont effacées (sauf si leur effacement contrevient à la contrainte 1b). L'ordre de cet effacement est fixe, cf. figure 6.9 :

- Les voyelles à gauche de la voyelle accentuée sont effacées en priorité par rapport aux voyelles à droite de la voyelle accentuée.
- Les voyelles les plus proches de la voyelle accentuée sont effacées en priorité par rapport à celles qui sont plus éloignées.
- Si une voyelle ne peut être effacée à cause d'autres contraintes plus forte, le processus d'effacement de voyelle est interrompu.

En ce qui concerne les contraintes sur la structure syllabique et les règles d'effacement, l'effacement d'une voyelle surnuméraire selon 3a peut être bloqué si son effacement a pour conséquence l'obtention d'une séquence \*ccc au sein d'une même syllabe. La priorité de 1b sur 3a explique l'existence de formes comme celle illustrée par *ħan.żér.na* « nous salissons » en 6.2<sup>34</sup>.

6.2 *ħan.żér.na* au lieu de \**ħmżér.na*

En 6.2 on s'attendrait en effet à ce que la première voyelle *a*, située à gauche de la seconde

34. Le placement de l'accent sur la seconde voyelle est la conséquence de la contrainte 4b rappelée ci-dessous.

voyelle (qui est accentuée), tombe en priorité. Mais sa disparition provoquerait l'apparition d'une séquence *fmz* et formerait ainsi une séquence \*CCC illicite en maltais. Le processus d'effacement est ainsi interrompu et la forme finale comporte trois syllabes au lieu des deux attendues.

**Contraintes d'accentuation en maltais :**

4a Une syllabe finale ne peut être accentuée que si elle est également l'unique syllabe du mot en question. Si d'autres syllabes la précèdent, leur voyelle est effacée selon les mêmes contraintes que pour la règle 3a.

4b Certains affixes provoquent le déplacement de l'accent.

- Les préfixes imperfectifs provoquent un déplacement de l'accent d'une syllabe vers la gauche.
- Les suffixes des premières et deuxièmes personnes du perfectif, commençant tous les quatre par une consonne, provoquent le déplacement de l'accent d'une syllabe vers la droite.
- Les suffixes commençant par une voyelle (le pluriel en *-u* et le *-et* du 3SG.F.IPFV) ne provoquent pas de déplacement d'accent.

4c Si la syllabe d'origine est plus lourde que la syllabe cible, l'accent ne se déplace pas, quel que soit l'affixe concerné.

4d Une voyelle longue est toujours accentuée.

Les contraintes d'accentuation sont hiérarchisées entre elles. En particulier, l'application de la règle 4b de déplacement de l'accent en fonction du type d'affixe avec lequel un radical se combine dépend de la contrainte 4c.

Par exemple, pour la forme *jmíss* « *il touche* » en 6.3, on s'attendrait à ce que la présence du préfixe de l'imperfectif *jv* provoque un déplacement de l'accent vers la gauche, en l'occurrence sur la voyelle du préfixe lui-même. Or le préfixe a une structure syllabique CV tandis que celle du radical est CVCC. La structure syllabique CVCC étant plus lourde que la structure CV, l'accent reste sur la syllabe la plus lourde. Une conséquence indirecte de cela est l'application de la règle 4a qui provoque la chute de la voyelle du préfixe et donne ainsi lieu à la forme *jmíss*.

6.3 *jmíss* au lieu de \**jímiss*

Par ailleurs, de même que les contraintes sur la structure syllabique, les contraintes

d'accentuation sont aussi soumises aux contraintes de suites de phonèmes licites. L'effacement de voyelles précédant une voyelle accentuée ne peut ainsi se faire s'il génère une séquence interdite par les contraintes 1a et 1b. Un exemple de cela est la forme *ħan.żirt* « je/tu salis » en 6.4.

6.4 *ħan.żirt* au lieu de *\*ħmżirt*

Dans le cas de 6.4, on s'attendrait à ce que la première voyelle tombe en vertu de la contrainte 4a. Mais cela induirait une séquence *ħmż*, or les séquences de trois consonnes à la suite sont illicites en maltais. La première voyelle se maintient donc.

En revanche, la contrainte 4d sur l'accentuation des voyelles longues est une contrainte absolue au même titre que les contraintes sur les suites licites de phonèmes. Cela a pour conséquence qu'il ne peut y avoir plus d'une voyelle longue par mot.

La figure 6.10 résume ensuite la hiérarchie existant entre les contraintes et règles phonologiques et morphologiques du maltais.

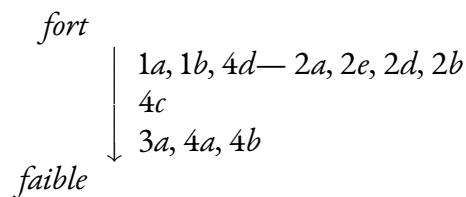


FIGURE 6.10 : Hiérarchie des contraintes et règles (mor-)phonologiques

#### 6.3.1.4 Réduction de la complexité de l'allomorphie radicale

Dans cette section, nous tirons les conséquences des observations qui précèdent pour l'analyse de l'allomorphie radicale en maltais et montrons comment la prise en compte des douze règles phonologiques et morphologiques ci-dessus permet de réduire la complexité de l'allomorphie radicale à une simple alternance vocalique affectant certaines classes de lexèmes.

Comme dans l'analyse de Camilleri (2011), nous utilisons la troisième personne du singulier masculin du perfectif comme forme de base unique dont seront dérivées les autres formes, notamment les autres radicaux. Cette forme a la caractéristique particulière de ne porter aucune marque flexionnelle affixale. Nous stipulons ainsi que la position de l'accent dans cette forme correspond à celle de l'accent lexical.

Ainsi que nous le montrons dans la suite de cette section, notre analyse diffère de celle de Camilleri en ce que nous ne stipulons pas de règles morphologiques donnant lieu à l'allomorphie de radical, notamment en termes de variation de la structure consonnes/voyelles, observable dans les paradigmes verbaux maltais. Dans l'analyse que nous proposons, toutes les formes du paradigme sont ainsi construites à partir de la forme de base. L'allomorphie radicale apparente n'est que le résultat d'ajustements phonologiques provoqués par le jeu des contraintes sur les extensions de la forme de base par les marques affixales que nous retrouvons dans les autres formes. Seuls les changements vocaliques restent morphologiquement déterminés.

Dans la suite de cette section, nous présentons ainsi les dérivations phonologiques mis en œuvre pour la construction des paradigmes des verbes à structure CVCVC, CVCCVC, CVCV, CVCC et enfin CV:C.

**6.3.1.4.1 Base verbale CVCVC** Les verbes à base verbale CVCVC comme KITEB « *écrire* » ou LAQAT « *frapper* » présentent une CSR de type ABCC DE, c'est-à-dire des paradigmes avec le maximum de radicaux différents pour cette base verbale. Par ailleurs, un verbe comme HATAF « *saisir* » ou SERAQ « *voler* » a la CSR ABCC AE et des verbes comme HARAQ « *brûler* » la CSR ABCC AC. Nous représentons dans le tableau 6.5 les paradigmes de ces cinq verbes avec leurs voyelles accentuées et leur découpage syllabique.

Les règles phonologiques et morphologiques décrites à la section précédente nous permettent de dériver les différentes structures syllabiques des radicaux, voire, lorsqu'aucun changement vocalique irrégulier n'est par ailleurs spécifié dans la composante morphologique, les formes des paradigmes complets de ces verbes.

Ainsi, pour le verbe KITEB « *écrire* » à CSR ABCC DE, la première personne du singulier du perfectif est obtenue par l'application du suffixe *-t* à la forme de base *ki.teb*<sup>35</sup>. Ce suffixe étant l'un de ceux qui provoquent un déplacement de l'accent (règle 4b), l'accent bascule sur la seconde et dernière syllabe (*kitébt*). Mais cette dernière ne peut être accentuée que si elle est l'unique syllabe du mot (contrainte 4a), ce qui a pour conséquence la suppression de la première voyelle et la resyllabification de l'ensemble en une seule et unique syllabe.

Par ailleurs, les suffixes provoquant le déplacement de l'accent sont aussi ceux qui entraînent un changement vocalique régulier *e* → *i* pour la deuxième voyelle de chaque forme. Le *e* de KÍ.TEB devient ainsi *i*, de sorte que la forme finale, après l'application de toutes les règles pertinentes et après changement vocalique, est *ktíbt*. Cette même dérivation

35. Nous donnons la forme de base, c'est-à-dire la troisième personne du singulier masculin du perfectif, avec l'accent lexical associé ainsi que son découpage syllabique.

		ABCC DE		ABCC AE		ABCC AC
		KITEB	LAQAT	HATAF	SERAQ	HARAQ
		« écrire »	« frapper »	« saisir »	« voler »	« brûler »
RAD2	PFV 1.SG	<i>ktíbt</i>	<i>lqátt</i>	<i>ḥtáf</i>	<i>sráqt</i>	<i>ḥráqt</i>
	PFV 2SG	<i>ktíbt</i>	<i>lqátt</i>	<i>ḥtáf</i>	<i>sráqt</i>	<i>ḥráqt</i>
	PFV 1PL	<i>ktíb.na</i>	<i>lqát.na</i>	<i>ḥtáf.na</i>	<i>sráq.na</i>	<i>ḥráq.na</i>
	PFV 2PL	<i>ktíb.tu</i>	<i>lqát.tu</i>	<i>ḥtáf.tu</i>	<i>sráq.tu</i>	<i>ḥráq.tu</i>
RAD1	PFV 3SG.M	<i>kí.teb</i>	<i>lá.qat</i>	<i>há.taf</i>	<i>sé.raq</i>	<i>há.raq</i>
RAD3	PFV 3SG.F	<i>kít.bet</i>	<i>lâq.tet</i>	<i>ḥát.fet</i>	<i>sér.qet</i>	<i>ḥár.qet</i>
RAD4	PFV 3PL	<i>kít.bu</i>	<i>lâq.tu</i>	<i>ḥát.fu</i>	<i>sér.qu</i>	<i>ḥár.qu</i>
RAD5	IPFV 1.SG	<i>ní.kteb</i>	<i>nó.lqot</i>	<i>ná.ḥtaf</i>	<i>nás.raq</i>	<i>náḥ.raq</i>
	IPFV 2SG	<i>tí.kteb</i>	<i>tó.lqot</i>	<i>tá.ḥtaf</i>	<i>tás.raq</i>	<i>táḥ.raq</i>
	IPFV 3SG.M	<i>jí.kteb</i>	<i>jó.lqot</i>	<i>já.ḥtaf</i>	<i>jás.raq</i>	<i>jáḥ.raq</i>
	IPFV 3SG.F	<i>tí.kteb</i>	<i>tó.lqot</i>	<i>tá.ḥtaf</i>	<i>tás.raq</i>	<i>táḥ.raq</i>
RAD6	IPFV 1PL	<i>níkt.bu</i>	<i>nól.qtu</i>	<i>náḥ.tfu</i>	<i>ná.sar.qu</i>	<i>ná.ḥar.qu</i>
	IPFV 2PL	<i>tíkt.bu</i>	<i>tól.qtu</i>	<i>táḥ.tfu</i>	<i>tá.sar.qu</i>	<i>tá.ḥar.qu</i>
	IPFV 3PL	<i>jíkt.bu</i>	<i>jól.qtu</i>	<i>jáḥ.tfu</i>	<i>já.sar.qu</i>	<i>já.ḥar.qu</i>

TABLE 6.5 : Paradigmes de verbes CVCVC

vaut également pour la deuxième personne du singulier de l'imperfectif et, au découpage syllabique près, pour les formes correspondantes du pluriel.

Pour les formes de la troisième personne du singulier féminin du perfectif et du pluriel perfectif, la dérivation est différente. Les suffixes *-et* et *-u* commençant par une voyelle ne sont pas des suffixes provoquant un déplacement de l'accent. Par ailleurs, ces suffixes comportent chacun une voyelle qui porte le nombre de syllabes dans les formes résultantes à trois. Afin de satisfaire la contrainte  $3a$  qui stipule que tout mot doit avoir deux syllabes, une des voyelles doit être effacée. Cet effacement se fait en priorité selon l'ordre indiqué à la figure 6.9 : en l'occurrence la voyelle effacée est la plus proche de la voyelle accentuée, à droite de cette dernière, c'est-à-dire la seconde voyelle de la forme de base. Les formes résultant de cette dérivation sont ainsi respectivement (après resyllabification) les formes *kít.bet* et *kít.bu*.

La dérivation complète pour toutes les formes du perfectif du paradigme de *KÍ.TEB* correspond ainsi à l'application des règles résumée à la figure 6.11. Cette dérivation vaut également pour tous les autres verbes à base verbale CVCVC.

Pour les formes de l'imperfectif, les préfixes de l'imperfectif provoquent un déplacement de l'accent de la forme de base sur la voyelle d'extension, cf. pour la première

	1SG	CV.CVC-t	$\xrightarrow{4b} \xrightarrow{4a}$	CCVct
RAD2	2SG	CV.CVC-t	$\xrightarrow{4b} \xrightarrow{4a}$	CCVct
	1PL	CV.CVC-na	$\xrightarrow{4b} \xrightarrow{3a}$	CCVC.na
	2PL	CV.CVC-tu	$\xrightarrow{4b} \xrightarrow{3a}$	CCVC.tu
RAD1	3SGM	CV.CVC	=	CV.CVC
RAD3	3SGF	CV.CVC-et	$\xrightarrow{3a}$	CVc.cet
RAD4	3PL	CV.CVC-u	$\xrightarrow{3a}$	CVc.cu

FIGURE 6.11 : Perfectif des verbes à base verbale CVCVC

personne du singulier et du pluriel (avec le  $-u$  du pluriel) :  $nV- + \text{KÍ.TEB} (+-u) \rightarrow \text{nikiteb}(u)^{36}$ . L'affixation produit également un nombre surnuméraire de syllabes dans le mot provoquant l'effacement des voyelles à droite de la voyelle thématique. Pour les formes du singulier, un seul effacement suffit pour satisfaire la contrainte  $3a$ . La forme obtenue est alors *ní.kteb*. Les formes du pluriel ont deux voyelles surnuméraires à effacer à cause de l'ajout syllabique dû au suffixe  $-u$  du pluriel. Les deux voyelles de la forme de base s'en trouvent ainsi effacées, de sorte que la forme dérivée pour la première personne du pluriel de l'imperfectif est *nik.tbu*.

Pour les autres verbes du tableau 6.5, les mêmes dérivations s'appliquent.  $\text{HÁ.TAF}$  ne se distingue de  $\text{KÍ.TEB}$  qu'en ce que les formes du radical RAD5 sont identiques à celles de RAD1 puisque la seconde voyelle  $a$  n'est affectée par aucun changement vocalique. La dérivation produit ainsi en surface la CSR  $\text{ABCC AE}$  observée par Camilleri (2011).

Pour le verbe  $\text{LÁ.QAT}$ , la morphologie impose un changement vocalique au singulier de l'imperfectif pour la seconde voyelle de la forme de base :  $a \rightarrow o$ . Ce changement vocalique a pour conséquence que la forme du perfectif singulier (à la première personne) devient *nó.lqot*, la voyelle thématique s'accordant avec celle du radical par harmonie vocalique<sup>37</sup>.

Le même phénomène explique également l'alternance  $e/a$  pour la première voyelle de  $\text{SÉ.RAQ}$  au radical RAD6. La morphologie impose un changement vocalique à la première voyelle de la forme de base au pluriel de l'imperfectif. Le reste de la dérivation est régulier.

36. La couleur de la voyelle thématique est ici  $i$  comme dans le cas régulier pour un mot commençant par une consonne non arrière. Nous revenons plus bas et surtout au chapitre 7 sur la sélection de la bonne voyelle thématique dans les cas plus compliqués.

37. Voir l'analyse de la flexion maltaise au chapitre 5, pages 329 à 338.

La dérivation des formes de l'imperfectif est résumée à la figure 6.12.

RAD5	1SG	nV-CV.CVC	$\xrightarrow{4b} \xrightarrow{3a}$	nVC.CVC
	2SG	tV-CV.CVC	$\xrightarrow{4b} \xrightarrow{3a}$	tVC.CVC
	3SGM	jV-CV.CVC	$\xrightarrow{4b} \xrightarrow{3a}$	jVC.CVC
	3SGF	tV-CV.CVC	$\xrightarrow{4b} \xrightarrow{3a}$	tVC.CVC
RAD6	1PL	nV-CV.CVC-u	$\xrightarrow{4b} \xrightarrow{3a}$	nVC.CCu
	2PL	tV-CV.CVC-u	$\xrightarrow{4b} \xrightarrow{3a}$	tVC.CCu
	3PL	jV-CV.CVC-u	$\xrightarrow{4b} \xrightarrow{3a}$	jVC.CCu

FIGURE 6.12 : Imperfectif des verbes à base verbale CVCVC

En revanche, les verbes SÉ.RAQ et HÁ.RAQ ayant pour seconde consonne une consonne sonante, ils ne se fléchissent pas pour le RAD6 en suivant la dérivation classique supprimant les deux voyelles du radical. Afin d'éviter l'émergence d'une séquence RR, certains dialectes du maltais, dont le MS, construisent par la morphologie des formes alternatives. Ces formes comportent les affixes et la voyelle d'extension réguliers du pluriel de l'imperfectif. Pour le cas de la première personne, nous aurons ainsi à gauche du radical le *na-* de la première personne (préfixe *n* et voyelle thématique *a*) provoquant un déplacement de l'accent lexical vers la gauche et le suffixe *-u* du pluriel à droite du radical. En revanche, pour le radical lui-même, la morphologie substituera un radical de l'imperfectif pluriel de structure CVCC au radical de base. Nous illustrons cette construction pour la première personne du pluriel *ná.sar.qu* du verbe SERAQ « voler » par la figure 6.13.

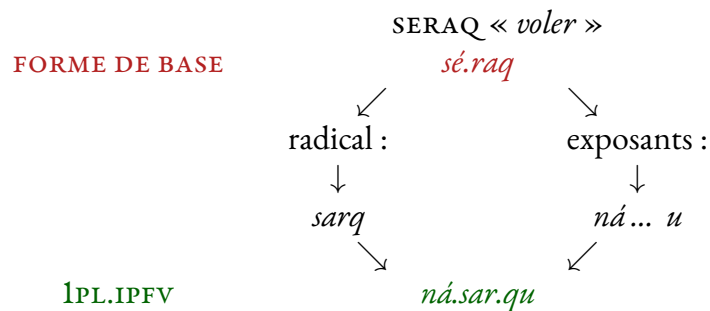


FIGURE 6.13 : RAD6 alternatif pour les verbes à base verbale CVRVC

6.3.1.4.2 Base verbale CVCCVC Pour les verbes à base verbale CVCCVC, les radicaux peuvent tous être obtenus de façon régulière. Compte tenu du nombre de consonnes, la règle 4a d’effacement des voyelles précédant la voyelle accentuée est bloquée par l’interdiction 1b d’avoir une suite de trois consonnes dans une même syllabe.

Nous donnons des exemples de verbes à base verbale CVCCVC au tableau 6.6.

		ABCC BC			BBCC BC	
		HANZER « salir »	BENĠEL « blesser »	FIXKEL « faire obstruction »	BARQAM « roucouler »	MEĠLAQ « mâcher bruyamment »
RAD2	PFV 1SG	<i>han.zirt</i>	<i>ben.ġilt</i>	<i>fix.kilt</i>	<i>bar.qámt</i>	<i>meċ.láqt</i>
	PFV 2S	<i>han.zirt</i>	<i>ben.ġilt</i>	<i>fix.kilt</i>	<i>bar.qámt</i>	<i>meċ.láqt</i>
	PFV 1PL	<i>han.zér.na</i>	<i>ben.ġél.na</i>	<i>fix.kél.na</i>	<i>bar.qám.na</i>	<i>meċ.láq.na</i>
	PFV 2PL	<i>han.zír.tu</i>	<i>ben.ġíl.tu</i>	<i>fix.kíl.tu</i>	<i>bar.qám.tu</i>	<i>meċ.láq.tu</i>
RAD1	PFV 3MS	<i>han.zer</i>	<i>ben.ġel</i>	<i>fix.kel</i>	<i>bár.qam</i>	<i>méc.laq</i>
RAD3	PFV 3FS	<i>han.zret</i>	<i>ben.ġlet</i>	<i>fix.klet</i>	<i>bár.qmet</i>	<i>méc.lqet</i>
RAD4	PFV 3PL	<i>han.zru</i>	<i>ben.ġlu</i>	<i>fix.klu</i>	<i>bár.qmu</i>	<i>méc.lqu</i>
RAD5	IPFV 1SG	<i>nhán.zer</i>	<i>nbén.ġel</i>	<i>nfix.kel</i>	<i>nbár.qam</i>	<i>nméc.laq</i>
	IPFV 2SG	<i>thán.zer</i>	<i>tbén.ġel</i>	<i>tfix.kel</i>	<i>tbár.qam</i>	<i>tméc.laq</i>
	IPFV 3MS	<i>jhán.zer</i>	<i>jbén.ġel</i>	<i>jfix.kel</i>	<i>jbár.qam</i>	<i>jméc.laq</i>
	IPFV 3FS	<i>thán.zer</i>	<i>tbén.ġel</i>	<i>tfix.kel</i>	<i>tbár.qam</i>	<i>tméc.laq</i>
RAD6	IPFV 1PL	<i>nhán.zru</i>	<i>nbén.ġlu</i>	<i>nfix.klu</i>	<i>nbár.qmu</i>	<i>nméc.lqu</i>
	IPFV 2PL	<i>thán.zru</i>	<i>tbén.ġlu</i>	<i>tfix.klu</i>	<i>tbár.qmu</i>	<i>tméc.lqu</i>
	IPFV 3PL	<i>jhán.zru</i>	<i>jbén.ġlu</i>	<i>jfix.klu</i>	<i>jbár.qmu</i>	<i>jméc.lqu</i>

TABLE 6.6 : Paradigmes de verbes à base verbale CVCCVC

À l’imperfectif, l’accent ne se déplace pas suite à la préfixation imperfective : la syllabe constituée par le préfixe de type CV est plus légère que la première syllabe de la base (de type CVC) qui porte l’accent lexical. L’accent reste donc, en vertu de la règle 4c sur la première syllabe de la base, de sorte que la voyelle thématique, en tant que voyelle précédant la voyelle accentuée, est la première à être effacée par la règle 3a. La dérivation complète de ce sous-paradigme est ainsi obtenue par l’application régulière des contraintes phonologiques et morphologiques indiquées dans la figure 6.15.

6.3.1.4.3 Base verbale CVCV Les verbes à base verbale CVCV ont la particularité de se terminer par une voyelle, ce qui peut, selon les suffixes avec lesquelles la base se combine, déclencher certaines des règles morphologiques 2a, 2e et 2d. Le tableau 6.7 donne trois exemples de paradigmes de verbes à structure CVCV.



	1SG	CVC.CVC-t	$\xrightarrow{4b} \xrightarrow{1b}$	CVC.CVCt
RAD2	2SG	CVC.CVC-t	$\xrightarrow{4b} \xrightarrow{1b}$	CVC.CVCt
	1PL	CVC.CVC-na	$\xrightarrow{4b} \xrightarrow{1b}$	CVC.CVC.na
	2PL	CVC.CVC-tu	$\xrightarrow{4b} \xrightarrow{1b}$	CVC.CVC.tu
RAD1	3SGM	CVC.CVC	=	CVC.CVC
RAD3	3SGF	CVC.CVC-et	$\xrightarrow{3a}$	CVC.CCet
RAD4	3PL	CVC.CVC-u	$\xrightarrow{3a}$	CVC.CCu

FIGURE 6.14 : Perfectif des verbes à base verbale CVCCVC

	1SG	nV-CVC.CVC	$\xrightarrow{4c} \xrightarrow{3a}$	nCVC.CVC
RAD5	2SG	tV-CVC.CVC	$\xrightarrow{4c} \xrightarrow{3a}$	tCVC.CVC
	3SGM	jV-CVC.CVC	$\xrightarrow{4c} \xrightarrow{3a}$	jCVC.CVC
	3SGF	tV-CVC.CVC	$\xrightarrow{4c} \xrightarrow{3a}$	tCVC.CVC
	1PL	nV-CVC.CVC-u	$\xrightarrow{4c} \xrightarrow{3a}$	nCVC.CCu
RAD6	2PL	tV-CVC.CVC-u	$\xrightarrow{4c} \xrightarrow{3a}$	tCVC.CCu
	3PL	jV-CVC.CVC-u	$\xrightarrow{4c} \xrightarrow{3a}$	jCVC.CCu

FIGURE 6.15 : Imperfectif des verbes à base verbale CVCCVC

Ainsi, au perfectif, les suffixes *-t*, *-na* et *-tu* déclenchent la règle *2a* qui indique qu'une voyelle à la frontière droite d'un morphe se palatalise si ce morphe est combiné à un morphe à sa droite commençant par une consonne :  $v\#C \rightarrow vjC$

Le suffixe *-et* déclenche quant à lui la règle *2e* provoquant l'allongement de la voyelle finale du morphe précédent. En l'occurrence, cela se traduit par l'allongement de toutes les voyelles :  $v_1\#v_{\neq u} \rightarrow v_1:$

Un cas particulier est celui où la voyelle finale du radical est *i*. Dans ce cas la règle *2e* provoque la fusion de *i* et *e* en la diphtongue *ie* :  $i\#e \rightarrow ie$

Le suffixe *-u* déclenche enfin la règle *2d* provoquant l'allongement de la voyelle finale du morphe précédant et l'ajustement de la voyelle *u* du suffixe en la semi-voyelle *w* correspondante :  $v\#_u \rightarrow v:w$

		ABCD DD QARA « écrire »	ABCD EF BENA « construire »	ABCD ED BEDA « commencer »
RAD2	PFV 1SG	<i>qrájt</i>	<i>bnéjt</i>	<i>bdéjt</i>
	PFV 2SG	<i>qrájt</i>	<i>bnéjt</i>	<i>bdéjt</i>
	PFV 1PL	<i>qráj.na</i>	<i>bnéj.na</i>	<i>bdéj.na</i>
	PFV 2PL	<i>qráj.tu</i>	<i>bnéj.tu</i>	<i>bdéj.tu</i>
RAD1	PFV 3MS	<i>qá.ra</i>	<i>bé.na</i>	<i>bé.da</i>
RAD3	PFV 3FS	<i>qrá:t</i>	<i>bníet</i>	<i>bdíet</i>
RAD4	PFV 3PL	<i>qrá:w</i>	<i>bné:w</i>	<i>bde:w</i>
RAD5	IPFV 1SG	<i>niq.ra</i>	<i>níb.ní</i>	<i>níb.da</i>
	IPFV 2SG	<i>tiq.ra</i>	<i>tíb.ní</i>	<i>tíb.da</i>
	IPFV 3MS	<i>jiq.ra</i>	<i>jíb.ní</i>	<i>jíb.da</i>
	IPFV 3FS	<i>tiq.ra</i>	<i>tíb.ní</i>	<i>tíb.da</i>
RAD6	IPFV 1PL	<i>niq.rá:w</i>	<i>níb.nu</i>	<i>níb.dé:w</i>
	IPFV 2PL	<i>tiq.rá:w</i>	<i>tíb.nu</i>	<i>tíb.dé:w</i>
	IPFV 3PL	<i>jiq.rá:w</i>	<i>jíb.nu</i>	<i>jíb.dé:w</i>

TABLE 6.7 : Paradigmes de verbes à base verbale CVCV

Le reste des dérivations correspond aux cas vus plus haut.

RAD2	1SG	CV.CV-t	$\xrightarrow{2a} \xrightarrow{4b} \xrightarrow{4a}$	CCéjt
	2SG	CV.CV-t	$\xrightarrow{2a} \xrightarrow{4b} \xrightarrow{4a}$	CCéjt
	1PL	CV.CV-na	$\xrightarrow{2a} \xrightarrow{4b} \xrightarrow{3a}$	CCéj.na
	2PL	CV.CV-tu	$\xrightarrow{2a} \xrightarrow{4b} \xrightarrow{3a}$	CCéj.tu
RAD1	3SGM	CV.CV	=	CV.CV
RAD3	3SGF	CV.CV-et	$\xrightarrow{2e} \xrightarrow{4c} \xrightarrow{4a}$	CCV:t
RAD4	3PL	CV.CV-u	$\xrightarrow{2d} \xrightarrow{4c} \xrightarrow{4a}$	CCV:w

FIGURE 6.16 : Perfectif des verbes à base verbale CVCV

Pour l'imperfectif, les dérivations sont régulières et indiquées à la figure 6.17.

Pour les verbes dont la seconde consonne est une sonante, les formes du pluriel ont toutefois une structure différente, morphologiquement conditionnée, à savoir NV.CR'V:w

RAD5	1SG	nV-CV.CV	$\xrightarrow{4b} \xrightarrow{3a}$	nV.CCV
	2SG	tV-CV.CV	$\xrightarrow{4b} \xrightarrow{3a}$	tV.CCV
	3SGM	jV-CV.CV	$\xrightarrow{4b} \xrightarrow{3a}$	jV.CCV
	3SGF	tV-CV.CV	$\xrightarrow{4b} \xrightarrow{3a}$	tV.CCV
RAD6	1PL	nV-CV.CV-u	$\xrightarrow{4b} \xrightarrow{3a}$	nV.CCu
	2PL	tV-CV.CV-u	$\xrightarrow{4b} \xrightarrow{3a}$	tV.CCu
	3PL	jV-CV.CV-u	$\xrightarrow{4b} \xrightarrow{3a}$	jV.CCu

FIGURE 6.17 : Imperfectif des verbes à base verbale CVCV

(ici pour la première personne du singulier).

6.3.1.4.4 Base verbale CVCC Le tableau 6.8 donne deux exemples de paradigmes de verbes à structure CVCC.

		ABBC DD	
		RASS	MESS
		« presser »	« toucher »
RAD2	PFV 1SG	<i>rasséjt</i>	<i>messéjt</i>
	PFV 2SG	<i>rasséjt</i>	<i>messéjt</i>
	PFV 1PL	<i>rasséj.na</i>	<i>messéj.na</i>
	PFV 2PL	<i>rasséj.tu</i>	<i>messéj.tu</i>
RAD1	PFV 3MS	<i>ráss</i>	<i>méss</i>
RAD3	PFV 3FS	<i>rás.set</i>	<i>més.set</i>
RAD4	PFV 3PL	<i>ras.sé:w</i>	<i>mes.sé:w</i>
RAD5	IPFV 1SG	<i>nróss</i>	<i>nmíss</i>
	IPFV 2SG	<i>tróss</i>	<i>tmíss</i>
	IPFV 3MS	<i>jróss</i>	<i>jmíss</i>
	IPFV 3FS	<i>tróss</i>	<i>tmíss</i>
RAD6	IPFV 1PL	<i>nrós.su</i>	<i>nmís.su</i>
	IPFV 2PL	<i>trós.su</i>	<i>tmís.su</i>
	IPFV 3PL	<i>jrós.su</i>	<i>jmís.su</i>

TABLE 6.8 : Paradigmes de verbes à base verbale CVCC

Les verbes à base verbale CVCC présentent la particularité de se terminer par deux consonnes. Ils déclenchent ainsi la règle morphologique 2*b* quand ils se combinent avec un suffixe commençant par une consonne (*t* et *na*). Cette règle stipule que lorsqu'un morphe se terminant par deux consonnes se combine avec un autre morphe à sa droite commençant lui-même par une consonne, la séquence *ej* est insérée pour éviter la suite de trois consonnes :  $CC\#C \rightarrow CCEjC$

Par ailleurs, les verbes de cette base sont particulièrement irréguliers en termes de changements vocaliques. Ils se caractérisent notamment par une insertion morphologique systématique d'une voyelle<sup>38</sup> à la fin de la forme de base à la troisième personne du pluriel du perfectif. Lorsque cette voyelle rencontre le suffixe du pluriel *u*, cela déclenche alors l'application de la règle 2*d* :  $v\#u \rightarrow vw$

La voyelle longue résultant de 2*e* provoque, en tant que voyelle longue, le déplacement de l'accent lexical, quand bien même le suffixe du pluriel ne l'aurait à lui seul pas provoqué (contrainte 4*d*). La voyelle précédant cette voyelle longue n'est cependant pas effacée en vertu de l'interdiction d'une séquence \*CCC au sein même d'une syllabe. La forme résultant est ainsi bisyllabique avec l'accent sur la seconde syllabe.

Nous résumons les dérivations menant aux structures des formes du perfectif à la figure 6.18.

	1SG	C'VCC-t	$\xrightarrow{2b} \xrightarrow{4b} \xrightarrow{1b}$	CV.CCéjt
RAD2	2SG	C'VCC-t	$\xrightarrow{2b} \xrightarrow{4b} \xrightarrow{1b}$	CV.CCéjt
	1PL	C'VCC-na	$\xrightarrow{2a} \xrightarrow{4b} \xrightarrow{3a}$	CV.CCéj.na
	2PL	C'VCC-tu	$\xrightarrow{2a} \xrightarrow{4b} \xrightarrow{3a}$	CV.CCéj.tu
RAD1	3SGM	C'VCC	=	C'VCC
RAD3	3SGF	C'VCC-et	=	C'VC.cet
RAD4	3PL	C'VCC-u	$\xrightarrow{2d} \xrightarrow{4d} \xrightarrow{1b}$	CVC.C'V:w

FIGURE 6.18 : Perfectif des verbes à base verbale CVCC

À l'imperfectif, le fait que la structure de la forme de base soit une syllabe lourde de type CVCC bloque le déplacement de l'accent sur la voyelle thématique de la marque de l'imperfectif. L'accent reste ainsi sur le radical et la voyelle thématique, précédant la

38. La couleur de la voyelle dépend des lexèmes concernés.

voyelle accentuée, tombe. Le reste de la dérivation est parfaitement régulier et résumé à la figure 6.19.

RAD5	1SG	nV-CVCC	$\xrightarrow{4c} \xrightarrow{4a}$	nCVCC
	2SG	tV-CVCC	$\xrightarrow{4c} \xrightarrow{4a}$	tCVCC
	3SGM	jV-CVCC	$\xrightarrow{4c} \xrightarrow{3a}$	jCVCC
	3SGF	tV-CVCC	$\xrightarrow{4c} \xrightarrow{4a}$	tCVCC
RAD6	1PL	nV-CVCC-u	$\xrightarrow{4c} \xrightarrow{3a}$	nCVc.cu
	2PL	tV-CVCC-u	$\xrightarrow{4c} \xrightarrow{3a}$	tCVc.cu
	3PL	jV-CVCC-u	$\xrightarrow{4c} \xrightarrow{3a}$	jCVc.cu

FIGURE 6.19 : Imperfectif des verbes à base verbale CVCC

**6.3.1.4.5 Base verbale CV:C** Outre le fait d'être des verbes dont la forme de base comporte une seule syllabe, cette syllabe étant lourde (comportant une voyelle longue), les verbes à base verbale CV:C se distinguent par leur nombreux changements vocaliques imposés par la morphologie. Prenons par exemple le verbe SA:R « *devenir* » dont nous rappelons le paradigme dans le tableau 6.9. Le *a:* de la forme de base est remplacé par *i* dans les formes associées au RAD1 et par *i:* dans les formes associées aux RAD5 et RAD6. La longueur de la voyelle varie donc d'une forme à l'autre.

Au perfectif, la forme n'ayant qu'une seule syllabe, aucun déplacement de l'accent ou effacement de voyelle n'intervient à la combinaison avec le suffixe *-t*. Avec les suffixes *-et* et *-u*, aucun changement n'est requis puisque ces suffixes n'entraînent pas de déplacement de l'accent et le radical combiné au suffixe constitue exactement deux syllabes avec l'accent sur la première. Avec *-na* enfin, l'accent ne se déplace pas plus, puisque la syllabe CVC du radical est plus lourde que celle du suffixe. L'ensemble faisant également deux syllabes dont la pénultième est accentuée, aucun changement supplémentaire n'est requis. La dérivation est résumée à la figure 6.20.

À l'imperfectif, la longueur de la voyelle fait que la voyelle portant l'accent lexical est plus lourde que la voyelle thématique. L'accent reste donc à son emplacement lexical et la voyelle thématique est effacée. Cet effacement est le seul ajustement phonologique nécessaire après l'affixation et le changement vocalique morphologique. La dérivation est résumée à la figure 6.21.

		ABBB CC			
		SA:R	QA:M	KIEM	MIET
		« devenir »	« se lever »	« mourir »	« mourir »
RAD2	PFV 1SG	<i>sirt</i>	<i>qomt</i>	<i>komt</i>	<i>mitt</i>
	PFV 2SG	<i>sirt</i>	<i>qomt</i>	<i>komt</i>	<i>mitt</i>
	PFV 1PL	<i>sir.na</i>	<i>qom.na</i>	<i>kom.na</i>	<i>mit.na</i>
	PFV 2PL	<i>sir.tu</i>	<i>qom.tu</i>	<i>kom.tu</i>	<i>mit.tu</i>
RAD1	PFV 3MS	<i>sa:r</i>	<i>qa:m</i>	<i>kiem</i>	<i>miet</i>
RAD3	PFV 3FS	<i>sa:.ret</i>	<i>qa:.met</i>	<i>kie.met</i>	<i>mie.tet</i>
RAD4	PFV 3PL	<i>sa:.ru</i>	<i>qa:.mu</i>	<i>kie.mu</i>	<i>mie.tu</i>
RAD5	IPFV 1SG	<i>nsi:r</i>	<i>nqu:m</i>	<i>nku:m</i>	<i>mmu:t</i>
	IPFV 2SG	<i>ssi:r</i>	<i>tqu:m</i>	<i>tku:m</i>	<i>tmu:t</i>
	IPFV 3MS	<i>jsi:r</i>	<i>jqu:m</i>	<i>jku:m</i>	<i>jmu:t</i>
	IPFV 3FS	<i>ssi:r</i>	<i>tqu:m</i>	<i>tku:m</i>	<i>tmu:t</i>
RAD6	IPFV 1PL	<i>nsi:.ru</i>	<i>nqu:.mu</i>	<i>nku:.mu</i>	<i>mmu:.tu</i>
	IPFV 2PL	<i>ssi:.ru</i>	<i>tqu:.mu</i>	<i>tku:.mu</i>	<i>tmu:.tu</i>
	IPFV 3PL	<i>jsi:.ru</i>	<i>jqu:.mu</i>	<i>jku:.mu</i>	<i>jmu:.tu</i>

TABLE 6.9 : Paradigmes de verbes à base CV:C

	1SG	CV:C-t	$\xrightarrow{4b}$	CVCt
RAD2	2SG	CV:C-t	$\xrightarrow{4b}$	CVCt
	1PL	CV:C-na	$\xrightarrow{4b}$	CVC.na
	2PL	CV:C-tu	$\xrightarrow{4b}$	CVC.tu
RAD1	3SGM	CV:C	=	CV:C
RAD3	3SGF	CV:C-et	=	CV:.Cet
RAD4	3PL	CV:C-u	=	CV:.Cu

FIGURE 6.20 : Perfectif des verbes à base verbale CV:C

### 6.3.2 Mesures de longueurs de description des deux analyses

Nous avons implémenté notre description en Alexina<sub>PARSL</sub> pour pouvoir la comparer à l'analyse implémentée présentée au chapitre 5. Dans la suite de ce chapitre, nous désignerons cette dernière par le terme « *description morphologique* » et nous ferons

RAD5	1SG	nV-CV:C	$\xrightarrow{4c} \xrightarrow{4a}$	nCV:C
	2SG	tV-CV:C	$\xrightarrow{4c} \xrightarrow{4a}$	tCV:C
	3SGM	jV-CV:C	$\xrightarrow{4c} \xrightarrow{3a}$	jCV:C
	3SGF	tV-CV:C	$\xrightarrow{4c} \xrightarrow{4a}$	tCV:C
RAD6	1PL	nV-CV:C-u	$\xrightarrow{4c} \xrightarrow{3a}$	nCV:.Cu
	2PL	tV-CV:C-u	$\xrightarrow{4c} \xrightarrow{3a}$	tCV:.Cu
	3PL	jV-CV:C-u	$\xrightarrow{4c} \xrightarrow{3a}$	jCV:.Cu

FIGURE 6.21 : Imperfectif des verbes à base verbale CV:C

référence à l'analyse que nous avons proposée dans la section précédente par le terme « *analyse morphologique* ». Notons que les deux analyses implémentent les mêmes 560 entrées lexicales et génèrent les mêmes formes par l'application des règles spécifiées dans leur grammaires respectives.

Nous avons appliqué notre mesure de compacité définie à la section 6.1 aux implémentations de nos deux analyses de la flexion du premier binyan du maltais. Les résultats des mesures sur nos deux analyses sont indiqués au tableau 6.10 et représentés sous forme de diagramme de Kiviat à la figure 6.22. Nous constatons que la description est à peu près de la même longueur dans la partie grammaire pour les deux analyses. En revanche, la longueur de description du lexique est significativement plus élevée dans la description morphologique que dans la description phonologique.

La figure illustre par ailleurs le fait que l'information n'est pas répartie de la même manière dans les deux descriptions. En particulier, l'information liée aux schèmes et à la structure est bien plus importante dans la description morphologique que dans la description morphonologique. En revanche, et comme on s'y attend, les résultats sont inversés pour ce qui concerne les règles morphonologiques.

Ce résultat est cependant attendu dans la mesure où la morphophonologie est par définition indépendante de sous-classes et concerne ainsi nécessairement l'intégralité des données lexicales. Contrairement à l'association des règles morphologiques à certaines entrées lexicales seulement, aucune spécification lexicale n'est ainsi nécessaire pour déclencher la réalisation de règles morphonologiques.

	DESCRIPTION MORPHOLOGIQUE	DESCRIPTION MORPHONOLOGIQUE
<b>LEXIQUE</b>	6 506	5 330
schèmes	769	21
variantes	5 282	4 722
supplétion	439	570
<b>GRAMMAIRE</b>	5 685	5 669
définition des cases	404	403
schèmes	1 821	681
structure	2 187	1 475
opérations	442	385
(mor-)phonologie	830	2 724
<b>TOTAL</b>	12 191	10 998

TABLE 6.10 : Résultats des mesures de compacité sur les descriptions du maltais

### 6.3.3 L'interface phonologie-morphologie et les cas de morphologie autonome

Comment choisir entre règles phonologiques et morphologiques? Nos deux implémentations montrent ainsi que deux types d'analyse très différents permettent de rendre compte de la flexion du premier binyan du maltais. La première, inspirée de (Camilleri, 2011), traite l'intégralité des variations par des règles morphologiques; la seconde propose un traitement de la plus grande partie des alternances des structures en phonologie. La question du choix de la meilleure des deux analyses se pose alors.

Une réponse évidente consiste à chercher l'analyse la plus économique.

Un premier point est que, toutes choses égales par ailleurs, stipuler des règles phonologiques ou morphologiques est nécessairement plus économique que traiter les mêmes données en morphologie dans la mesure où les règles phonologiques et morphologiques une fois stipulées s'appliquent par définition à l'intégralité du système de la langue. Les règles morphologiques sont en revanche spécifiques à des classes et sous-classes du lexique; spécifier une règle morphologique pour une classe ne dispense



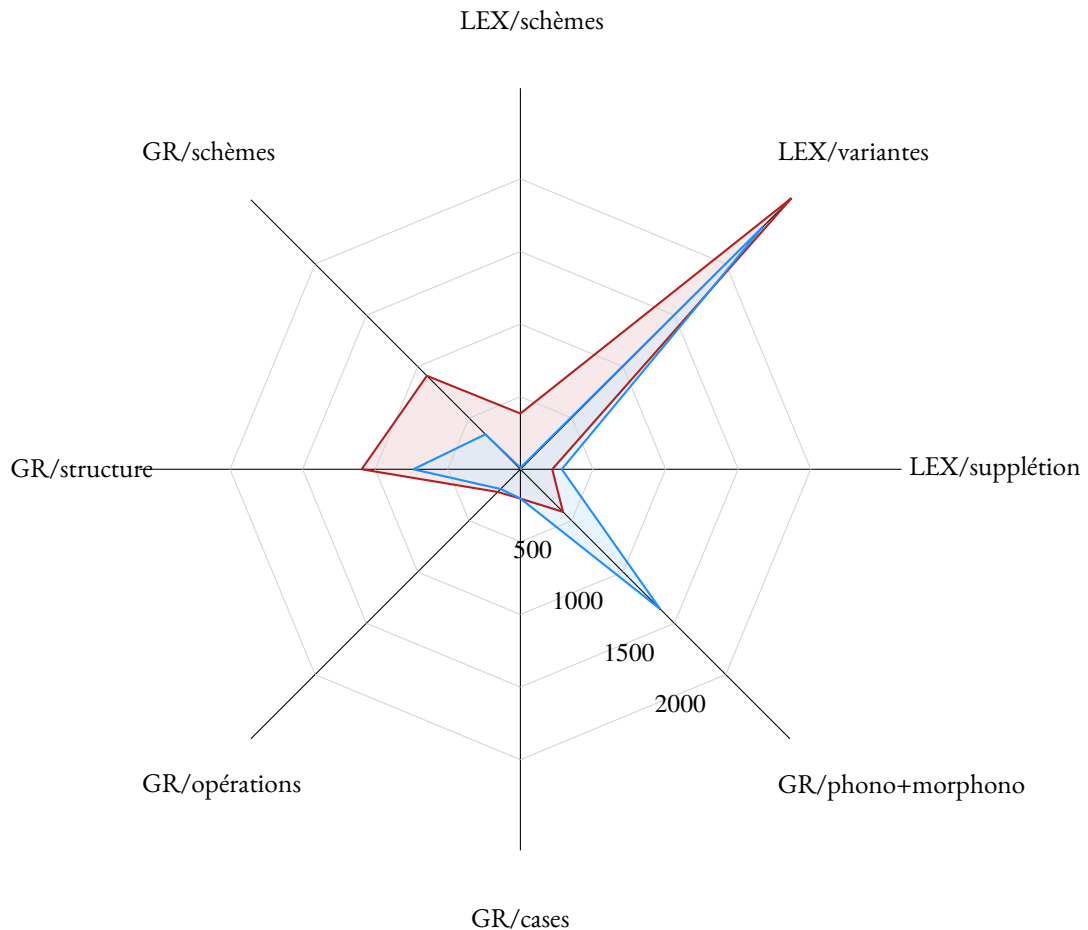


FIGURE 6.22 : Représentation des mesures de compacité sur les descriptions du maltais

pas du fait d'en stipuler pour d'autres. En l'occurrence, nos données du maltais peuvent être traitées de façon purement morphologique ou de façon mixte en introduisant une composante morphologique importante. Il va de soi que le traitement (mor-)phonologique n'a d'intérêt à cet égard que si les mêmes règles phonologiques s'avèrent avoir un effet sur la structure d'autres composantes de la langue, de sorte que la solution à des problèmes différents peut être factorisée en ces seules règles. Autrement, il n'y a plus de différence à stipuler des règles phonologiques plutôt que morphologiques en termes d'économie du système descriptif.

La question de l'économie maximale peut cependant être tranchée par l'observation des résultats de mesures externes. D'après la mesure de compacité que nous avons définie, les deux analyses ont la même longueur si l'on n'observe que la partie grammaticale. En revanche, l'analyse morphologique requiert significativement moins d'information dans la partie lexicale et est donc, en fin de compte, plus compacte.

Compte tenu des mesures indiquées à la section 6.3.2 et des considérations sur la factorisation de l'information en morphologie, nous avons ainsi objectivement intérêt à choisir la seconde description (ayant plus amplement recours à la morphologie) comme la description la plus compacte.

**Les cas de morphologie irréductibles** Cependant, au delà des questions d'économie descriptive, des problèmes supplémentaires, intervenant dans la construction des formes, subsistent qui ne peuvent être réduits à des règles phonologiques générales. Ces règles constituent de la morphologie pure au sens d'Aronoff (1994).

Les alternances vocaliques observables qui se superposent à la variation de la structure des radicaux ne sont pas réductibles à des alternances phonologiques régulières. Elles nécessitent un traitement spécifique ne s'appliquant qu'à certaines sous-classes de lexèmes. Cette restriction indique qu'elles ne peuvent relever de règles phonologiques générales et doivent donc appartenir à la composante morphologique. N'étant conditionnées ni par la phonologie, ni d'ailleurs par la syntaxe ou la sémantique, elles constituent une instance de morphologie pure, c'est-à-dire autonome.

Les règles phonologiques ne permettent également que partiellement d'identifier la voyelle d'extension correcte qui apparaît entre le radical et les préfixes de l'imperfectif. La voyelle en question n'est pas la même pour tous les verbes : *a* pour HATAF, *i* pour KITEB et BENA et *o* pour LAQAT. La couleur de la voyelle thématique dépend largement de la consonne initiale et de la première voyelle du radical de la forme de citation. Plus précisément, la voyelle est en général un *i*, mais elle s'ajuste par harmonie vocalique à la première voyelle du radical lorsque la consonne initiale de ce dernier est une consonne arrière (B : *għ, ħ, h, k, q*). Dans certains cas cependant, la couleur de la voyelle thématique ne correspond pas à ce qui est prévu par cette règle. Elle doit alors être stipulée indépendamment de la phonologie pour les lexèmes concernés. Il subsiste ainsi des cas irréductibles. Ces cas sont de la morphologie autonome.

## 6.4 Approche syntaxique ou morphologique ?

Dans cette section, nous étudions de plus près la répartition de l'information entre les composantes morphologique et syntaxique. De même que la phonologie et la morphologie, la syntaxe se distingue de la morphologie au sens strict par le fait que les règles qu'elle comporte s'appliquent indifféremment à tous les mots de la langue. Elles ne sont pas autant soumises à des idiosyncrasies lexicales similaires à celles qui gouvernent

l'appartenance à une catégorie flexionnelle, l'appartenance à une classe flexionnelle ou un type de défektivité. Ainsi, la syntaxe est, par le pouvoir factorisant systématique de ses règles, intrinsèquement plus compacte que la morphologie. Cela a pour conséquence que si un phénomène peut être décrit autant par des mécanismes syntaxiques que morphologiques, le décrire en syntaxe mènera à une description globale plus compacte. Cette question se pose de façon importante dans le cas des périphrases. Si une partie d'un paradigme peut être décrite comme mettant en œuvre de la périphrase, alors cette dernière peut être prise en compte par la syntaxe, allégeant la part de morphologie dans la grammaire globale et optimisant par là sa compacité.

Toutefois, s'il est par défaut plus économique de traiter des phénomènes par de la périphrase lorsque cela est possible, le bénéfice d'un arbitrage en faveur de la syntaxe est nul s'il impose par ailleurs un traitement complémentaire en morphologie. Dans cette section nous montrons par conséquent, à l'aide d'exemples de la flexion du kurde sorani, qu'il n'est pas toujours optimal de traiter des cas de périphrase potentielle en syntaxe. L'exemple des MARQUES PERSONNELLES MOBILES (MPM) (Samvelian, 2007) montre en effet qu'il peut parfois être préférable de maintenir une analyse synthétique pour des données *a priori* potentiellement périphrastiques.

Cette analyse repose sur les données et la première analyse de Samvelian (2007) que nous avons complétée par des données de descriptions existantes et de données originales récoltées auprès de locuteurs natifs de la langue (Walther, 2012a)<sup>39</sup>.

### 6.4.1 La flexion verbale et le marquage personnel en kurde sorani

Le kurde sorani est une langue iranienne occidentale parlée essentiellement à la frontière irano-irakienne. Il existe une variante standard de la langue qui constitue également le fondement du sorani écrit. Ce standard correspond essentiellement aux dialectes parlés dans les villes de Sulaymaniye et d'Arbil en Irak. Dans la suite de ce travail nous employons le terme générique SORANI pour désigner ce dialecte standard. Il correspond à ce que Haig (2010) qualifie de *suleimani*, dialecte kurde central.

---

39. Dans l'article original traitant de ces données (Walther, 2012a), nous avons montré que pour rendre compte du comportement des clitiques du sorani il est nécessaire d'admettre une part de structure construite en morphologie qui serait accessible à la syntaxe. En particulier, le comportement erratique de ces clitiques au parfait ne peut s'expliquer sans faire l'hypothèse d'une structure interne dans le mot-forme, réalisée en morphologie et accédée en syntaxe. Nous avons montré qu'il n'est pas nécessaire dans une approche lexicaliste de souscrire à un rejet global de toute structure en morphologie. Notre analyse proposait une solution dans laquelle la morphologie fournit à la syntaxe non pas des éléments constituant des blocs bruts et dans lesquels aucune structure n'est accessible, mais une forme structurée laissant apparaître un *relief* interne transparent en syntaxe.

Le sorani se distingue des autres variantes du kurde par sa morphologie verbale complexe et surtout par son système compliqué d'ENDOCLITIQUES<sup>40</sup> aussi qualifiés de MARQUES PERSONNELLES MOBILES (*mobile person markers*) par Bonami & Samvelian (2008) qui font l'objet de cette section. Dans la suite nous nous référons à ces clitiques par l'acronyme MPM.

Dans notre travail, nous utilisons une transcription du sorani qui s'appuie sur les transcriptions habituellement employées pour l'étude des langues iraniennes<sup>41</sup>. Elle est indiquée au tableau D.1 de l'annexe D.

### 6.4.1.1 Composition générale des formes verbales

Les marques flexionnelles verbales en sorani sont constituées de préfixes et de suffixes marquant essentiellement le temps, l'aspect, le mode et la polarité pour les premiers et la personne (et parfois les TAM) pour les seconds. La distribution des affixes peut être décrite par des classes de position que nous présentons au tableau 6.11. Y figurent dans la position -2 les marques de négation *ma-* et *na-*, en -1 les marques de mode *da-* (indicatif) et *bi* (subjunctif), en 0 les radicaux, dans la classe grisée *P* les marques du passif<sup>42</sup>, dans la classe 1 les marques du parfait (dont le subjunctif parfait *b*), en 2 les différentes marques personnelles standard et enfin en 3 la marque du subjunctif parfait *ayê*.

-2	-1	0	(P)	1	2	3
ma	da	RADICAL	râ	+û	im	ayê
na	bi	(1, 2, 3, 4)	rê	+bû	î(t)	
nâ				b	ê(t)	
					a	
					în	
					in	
					Ø	
				n		

TABLE 6.11 : Classes de positions affixales dans une analyse PFM du système verbal sorani

40. Clitiques qui s'attachent à l'intérieur de leur hôte.

41. Voir par exemple (Lazard et al., 2006).

42. Nous avons indiqué le placement des marques du passif dans le tableau 6.11. Toutefois, dans l'analyse que nous avons proposée du passif en sorani dans (Walther, 2011a), nous ne les considérons pas comme des marques flexionnelles mais comme des suffixes dérivationnels. Cette distinction n'étant cependant pas essentielle pour la question du traitement périphrastique ou synthétique du parfait en sorani, nous ne l'approfondissons pas plus ici.

## 6.4.1.2 Marquage de la personne

Les désinences personnelles du sorani sont habituellement suffixées. Elles se classent en trois paradigmes : deux paradigmes de désinences verbales au sens propre (D1 et D2) et un paradigme formé par les formes enclitiques du verbe *bun* « être » (D3). Les D1 sont employées pour les formes du présent et le subjonctif du parfait construit à partir du RAD1. Les D2 servent à réaliser les formes de l'imparfait, du prétérit et du plus-que-parfait. Enfin, les D3 sont employées pour les formes restantes. À ces désinences personnelles s'ajoutent les marques personnelles mobiles (MPM) décrites ci-dessous.

Nous présentons les quatre paradigmes de marques flexionnelles du sorani au tableau 6.12.

Marques personnelles		D3	Marques personnelles enclitiques
D1	Marques personnelles du présent		
	<i>-m</i>		<i>-în</i>
	<i>-î(t)</i>		<i>-n</i>
	<i>-ê(t)</i>		<i>-n</i>
D2	Marques personnelles du passé	MPM	Marques personnelles mobiles
	<i>-m</i>		<i>-mân</i>
	<i>-î(t)</i>		<i>-tân</i>
	<i>-ø</i>		<i>-yân</i>

TABLE 6.12 : Marquage de la personne en sorani

## 6.4.1.3 Propriétés des MPM

Les marques personnelles mobiles (MPM) remplissent, selon les cas, le rôle de marqueur de sujet ou d'objet. Aux temps du présent, le marquage du sujet est assuré par les désinences personnelles normales ; l'objet peut être marqué de façon facultative par les MPM, cf. l'exemple 6.5. En revanche, au passé, nous avons à faire à une INVERSION MORPHOLOGIQUE (*morphological reversal*) au sens de Baerman (2007a). Ce sont les MPM qui marquent le sujet, cf. l'exemple 6.6. Le marquage de l'objet est alors pris en charge par les désinences normales ; il est optionnel dès lors qu'il existe déjà un syntagme nominal dans la phrase ayant la fonction d'objet. L'inversion morphologique entre le présent et le passé dans le marquage du sujet et de l'objet grâce aux MPM ou aux désinences normales est analysée par Haig (2010) comme une TRACE D'ERGATIVITÉ SCINDÉE, disparue du système nominal en sorani mais encore visible dans les autres variantes du kurde<sup>43</sup>

43. Nombre de variantes du kurde du nord, comme le kurmandji, ainsi que d'autres dialectes kurdes centraux plus proches comme le mukri (Mann, 1906) ils affichent en effet toujours une opposition entre

6.5 nâ=yân=xo-n  
 NEG=OBJ.3PL=manger.RAD1-3PL  
 « Ils ne les mangent pas. »

6.6 na=yân=xward-in  
 NEG=SUBJ.1PL=manger.RAD2-3PL  
 « Ils ne les ont pas mangé(e)s. »

Mais la caractéristique la plus marquante des MPM est que, contrairement aux autres marques, elles ne peuvent par ailleurs pas être associées à une classe de position unique. Lorsqu'elles apparaissent sur une forme verbale, elles peuvent occuper jusqu'à cinq positions différentes. Les positions d'insertion possibles pour les MPM sont marquées par les créneaux d'insertion notés I<sub>1</sub> à I<sub>5</sub> au tableau 6.13<sup>44</sup>.

-2	I <sub>1</sub>	-1	I <sub>2</sub>	0	I <sub>3</sub>	1	I <sub>4</sub>	2	I <sub>5</sub>	3
ma		da		RADICAL (1, 2, 3, 4)		+û		im		ayê
na		bi				+bû		î(t)		
nâ						b		ê(t)		
								a		
								în		
								in		
								Ø		
								n		

TABLE 6.13 : Positions possibles pour les MPM à l'intérieur d'un verbe transitif

Nous donnons ci-dessous des exemples pour chacune des cinq possibilités de placement. La MPM s'y attache respectivement après le radical, en position I<sub>1</sub> après le préfixe négatif pour l'exemple 6.7, en position I<sub>2</sub> après le préfixe de l'indicatif en 6.8 et après le préfixe négatif porte-manteau du présent en 6.9 et enfin en position I<sub>3</sub> après le radical pour l'exemple 6.10 (verbe sans marque préfixale).

6.7 na=tân=-da-xward-in  
 NEG=2PL=IND-manger.RAD2-3PL  
 « Vous ne les mangiez pas. »

cas direct et oblique.

44. Ce tableau ne contient pas de classe de position réservée au passif, un les MPM ne s'attachant qu'à des verbes transitifs auxquels n'appartiennent évidemment pas les verbes passifs.

6.8 da=yân=xo-n  
 IND=3PL=manger.RAD1-2PL  
 « Vous les mangez. »

6.9 nâ=yân=xo-n  
 NEG.PRS=3PL=manger.RAD1-2PL  
 « Vous ne les mangez pas. »

6.10 xward=mân=in  
 manger.RAD2=1PL=3PL  
 « Nous les avons mangé(e)s. »

Par ailleurs, les MPM peuvent également se retrouver en position I<sub>4</sub> et I<sub>5</sub> lorsque le verbe est au parfait (marqué par -û) ou au plus-que-parfait (marqué par -bû). Ainsi, dans les exemples 6.11 à 6.14, la MPM apparaît-elle respectivement en position I<sub>4</sub> pour les exemples 6.11 et 6.12 et I<sub>5</sub> pour les exemples 6.13 à 6.14.

6.11 xward-û=mân=in  
 manger.RAD2-û.PRS.PRF=1PL=3PL  
 « Nous les avons mangé(e)s. »

6.12 xward-bû=mân=in  
 manger.RAD2-bû.PST.PRF=1PL=3PL  
 « Nous les avions mangé(e)s. »

6.13 xward-û-in=î  
 manger.RAD2-ûPRS.PRF-3PL=3SG  
 « Il les a mangé(e)s. »

6.14 xward-bû-in=î  
 manger.RAD2-bû.PST.PRF-3PL=3SG  
 « Il les avait mangé(e)s. »

### 6.4.2 Analyse du placement des MPM pour les formes n'exprimant pas le parfait

Une première description détaillée des propriétés de placement des MPM en kurde sorani a été donnée par Samvelian (2007) pour les formes n'exprimant pas le parfait. Notons cependant au préalable que comme l'indique Samvelian (2007) et comme l'illustrent les données de McCarus (1958), le placement des MPM n'est pas conditionné par la phonologie. Il n'est notamment pas possible de déduire le placement des MPM du placement de l'accent. Ainsi, contrairement au comportement des clitiques du pashto décrit par Anderson (2005), les MPM du sorani affichent des propriétés de placement déterminées par la seule morphologie. Une simplification par la phonologie de ces données (comme celle opérée à la section 6.3 de ce chapitre) ne peut ainsi être réalisée.

Samvelian (2007) indique dans son analyse que les MPM clitiques se placent systématiquement en seconde position, qu'elle soit interne ou externe. Les règles de placement énoncées par Samvelian sont les suivantes :

- Si le verbe n'est pas l'élément initial du SV, la MPM se rattache comme un enclitique au premier élément du SV. Il est alors un enclitique de seconde position.
- Si le verbe est en position initiale dans le SV, la MPM est un endoclitique de seconde position dans le verbe.

Les MPM peuvent ainsi, selon les cas, être des ENCLITIQUES ou des ENDOCLITIQUES<sup>45</sup>. Pour les MPM endoclitiques, la règle générale (RG) spécifie ainsi un placement en seconde position interne. La règle générale de seconde position interne est illustrée par les exemples 6.7 à 6.10 ci-dessus.

**RG (règle générale) :** lorsqu'un verbe est en position initiale dans le SV, la MPM est un endoclitique qui se place en seconde position, c'est-à-dire après le premier morphe.

---

45. Certains auteurs préfèrent le terme de MÉSACLISE pour les cas où un clitique s'insère entre des morphes identifiables et réservent le terme d'endoclise au cas où le clitique s'insère à l'intérieur du radical. Un exemple de cet emploi est l'étude de Crismann (2002) sur la mésoclise en portugais d'Europe. L'endoclise à l'intérieur même du radical est traitée dans le travail de Harris (2002) pour les verbes udi : plus précisément, les clitiques de l'udi affichent à la fois des cas d'intertion inter-morphe et à l'intérieur du radical verbal. Harris emploie néanmoins le terme d'endoclise pour les deux phénomènes. Dans la suite de ce travail, nous suivons Harris en employant le terme d'endoclise comme un terme générique qualifiant le rattachement d'un clitique au sein de la structure phonologique constituée par un mot donné. L'intégrité phonologique du mot en question se trouve ainsi brisée par l'insertion clitique. Si l'on cherchait à distinguer les deux notions, le terme de mésoclise serait néanmoins à préférer pour décrire le placement des MPM en sorani.



Plus précisément, le placement en seconde position interne selon la règle générale est réalisé comme suit <sup>46</sup> :

1. Quand un MPM est un endoclitique, elle s'insère par défaut en deuxième position dans le verbe, c'est-à-dire après le premier préfixe.
  - a) Le premier préfixe peut être le préfixe de l'indicatif présent et imparfait *da-*,
  - b) le préfixe du subjonctif *bi-*,
  - c) ou encore un des préfixes négatifs *na-*, *nâ-* ou *ma-*.

Dans ces cas, la MPM s'insère ainsi en I<sub>1</sub> pour en présence d'un des suffixes *na-* et *ma-*, ou en I<sub>2</sub> après les préfixes *da-* et *bi-* si la position -2 n'est pas remplie. En présence du préfixe négatif portemanteau *nâ-*, la MPM s'insère également par défaut en position I<sub>2</sub>.

2. S'il n'y a pas de préfixe, la MPM s'insère par défaut immédiatement après le radical, c'est-à-dire en position I<sub>3</sub>.

En ce qui concerne le placement interne, Samvelian (2007) précise cependant qu'il ne se résume pas toujours à un simple placement après le premier préfixe ou, en cas d'absence de préfixe, après le radical. Elle donne des exemples montrant que le placement des MPM peut parfois être soumis à une contrainte supplémentaire, plus spécifique, exception à la règle générale. Cette contrainte E-3SG régit notamment le placement des MPM dans les exemples 6.15 and 6.16 ci-dessous.

**E-3SG (première exception) :** La marque du sujet à la troisième personne du singulier, qu'elle soit selon les temps désinence régulière ou MPM, se place toujours après le marqueur objet.

6.15    xward-in=**î**  
          manger.RAD2-3PL=**3SG**  
          « *Il les a mangé(e)s.* »

6.16    na-xward-in=**î**  
          NEG-manger.RAD2-3PL=**3SG**  
          « *Il ne les a pas mangé(e)s.* »

---

46. Les positions indiquées font référence aux colonnes grisées dans le tableau 6.13.

En plus de la contrainte spécifique mentionnée par Samvelian (2007), McCarus (1958) note qu'il en existe une autre qui, à sa manière, interfère avec la règle générale 6.4.2<sup>47</sup>.

**E-1SG (seconde exception) :** Qu'elle marque le sujet ou l'objet, une marque personnelle de la première personne du singulier précède nécessairement toute marque personnelle du pluriel.

Nous donnons des exemples de la règle E-1SG ci-dessous. Les formes dans ces exemples ne s'expliquent pas sans la règle E-1SG.

6.17 xward=**m**=in  
manger.RAD2=**1SG**=3PL  
« *Je les mangeai.* »

6.18 xward-im=**tân**  
manger.RAD2-1SG=**2PL**  
« *Vous me mangeâtes.* »

6.19 na=**m**=xward-in  
NEG=**1PL**=manger.RAD2-3PL  
« *Je ne les mangeai pas.* »

6.20 na-xward-im=**tân**  
NEG-manger.RAD2-1SG=**2PL**  
« *Vous ne me mangeâtes pas.* »

En plus de la règle générale de placement en seconde position, deux contraintes supplémentaires déterminent donc les propriétés de placement des MPM. Ces deux contraintes peuvent forcer l'insertion de la MPM après la désinence personnelle normale. Les désinences personnelles étant associées à la classe de position 2, ces règles supplémentaires peuvent ainsi provoquer l'insertion de la MPM en position I<sub>5</sub>.

Les contraintes formulées par Samvelian (2007) sur le placement interne des MPM ont subséquemment été formalisées dans (Bonami & Samvelian, 2008). Bonami & Samvelian formalisent les contraintes de placement des clitiques du sorani décrites par Samvelian

47. Les deux exceptions à la règle de placement générale sont présentes dans (McCarus, 1958). Samvelian (2007) ne mentionne que l'exception E-3SG.

(2007) pour les MPM au présent et au passé non parfait<sup>48</sup>. La formalisation des propriétés morphologiques de placement par Bonami & Samvelian emploie des concepts introduits par Crysmann (2002) pour le traitement de l'insertion dans la forme verbale des clitiques personnels du portugais d'Europe. L'approche de Crysmann repose sur l'idée que la morphologie manipule des éléments qui sont des LISTES DE MORPHES structurées (*morph lists*) et non des segments de phonologie bruts. Elles contiennent des éléments qui ne sont ni des morphes indépendants, ni des assemblages de morphes opaques. Les éléments opaques de la liste de morphes sont également appelés RADICAUX MORPHOLOGIQUES (*morphological stems*) par les auteurs. Nous donnons un exemple de la représentation d'une telle liste de morphes à deux éléments en 6.21.

6.21    nâ-xo-m  
 NEG-manger.RAD1-1SG  

$$\left[ \begin{array}{l} \text{MORPHES } \langle n\hat{a}, xo, m \rangle \\ \text{TRAITS } \left[ \begin{array}{l} \text{ASPECT } iprf \\ \text{TEMPS } prs \\ \text{POLARITÉ } \textit{négative} \end{array} \right] \end{array} \right]$$
 « Je mange. »

Dans (Bonami & Samvelian, 2008), la réalisation des formes rendant compte de l'insertion des éléments est représentée par l'application successive des règles de réalisation bloc par bloc. Suite à l'application des règles d'un bloc, une opération de compactage peut générer des listes de morphes opaques au sens de Crysmann (2002). L'opération en question est appelée *compact*. Elle est déclenchée lorsqu'un verbe au passé comporte un sujet à la troisième personne du singulier. En ce cas, le radical et tous les préfixes de TAM ou de négation qui le précèdent sont compactés avec la marque personnelle normale D2 du sujet en un seul et unique élément de la liste de morphe, cf. l'exemple 6.22.

6.22

48. Bonami & Samvelian (2008) formalisent également le placement des MPM à l'impératif. Dans la suite, nous n'explicitons pas le comportement des MPM à l'impératif, ce comportement étant parfaitement pris en charge par les règles posées par Bonami & Samvelian.

$$\text{compact} \left( \left[ \begin{array}{l} \text{MORPHES } \langle xward, in \rangle \\ \text{TRAITS } \left[ \begin{array}{l} \text{ASPECT } iprf \\ \text{TEMPS } \textit{prétérit} \\ \text{POLARITÉ } \textit{positive} \end{array} \right] \\ \text{SUJET } \left[ \begin{array}{l} \text{PERSONNE } 3 \\ \text{NOMBRE } sg \end{array} \right] \end{array} \right] \right) = \left[ \begin{array}{l} \text{MORPHES } \langle xwardin \rangle \\ \text{TRAITS } \left[ \begin{array}{l} \text{ASPECT } iprf \\ \text{TENSE } \textit{prétérit} \\ \text{POLARITÉ } \textit{positive} \end{array} \right] \\ \text{SUJET } \left[ \begin{array}{l} \text{PERSON } 3 \\ \text{NOMBRE } sg \end{array} \right] \end{array} \right]$$

La MPM objet étant introduite par un bloc ultérieur, elle ne peut plus s'insérer à l'intérieur de l'élément compacté et la seconde position accessible se situe après l'élément compacté, c'est-à-dire en position I<sub>5</sub> après la marque personnelle D2 selon le tableau 6.13, cf. l'exemple 6.23. L'opération de compactage permet ainsi de rendre compte des règles de placement exceptionnelles pour les MPM.

6.23 xward-in=î

manger.RAD2-3PL=3SG

$$\left[ \begin{array}{l} \text{MORPHES } \langle xwardin \rangle \oplus \langle \hat{i} \rangle \\ \text{TRAITS } \left[ \begin{array}{l} \text{ASPECT } iprf \\ \text{TEMPS } \textit{prétérit} \\ \text{POLARITÉ } \textit{positive} \end{array} \right] \\ \text{SUJET } \left[ \begin{array}{l} \text{PERSONNE } 3 \\ \text{NOMBE } sg \end{array} \right] \end{array} \right]$$

« Il les a mangé(e)s. »

Bonami & Samvelian (2008) formalisent le placement des MPM tel qu'il est décrit par Samvelian (2007). Or, en plus de la règle générale RG et de la première exception E-3SG à cette règle également notée par Samvelian, McCarus (1958) avait noté une seconde exception E-1SG à règle générale RG. Cette règle, omise par Samvelian (2007), manque également dans la formalisation des MPM par Bonami & Samvelian (2008).

Afin de satisfaire la contrainte sur les marques de la première personne du singulier, il suffit cependant simplement d'ajouter les deux règles suivantes provoquant la même

opération de compactage<sup>49</sup> :

1. Quand une phrase comporte

- un verbe au présent et un sujet à la première personne du singulier et un objet au pluriel,
- ou un verbe au passé et un objet à la première personne du singulier et un sujet au pluriel,

une autre opération `compact` s'applique. Elle suffit à rendre compte de la séquence opaque constituée du radical et la désinence personnelle à la première personne du singulier. La MPM s'attache à l'extérieur de cette séquence opaque.

### 6.4.3 Prise en compte des formes du parfait

#### 6.4.3.1 Les données manquantes

Grâce au mécanisme de compactage, Bonami & Samvelian (2008) parviennent à formaliser le comportement de MPM pour les verbes n'exprimant pas le parfait comme un placement systématique en deuxième position, qu'elle soit interne (c'est-à-dire au sein de la liste de morphes) ou externe (comme enclitique du premier mot du SV). Ils font par ailleurs le choix explicite de laisser de côté les formes exprimant le parfait. Suite à l'extension de leur formalisation par la prise en compte de la contrainte E-1SG que nous avons proposée dans la section précédente, nous constatons que leur formalisation permet effectivement de rendre compte de l'intégralité des formes qu'ils étudient.

Cependant, si l'on prend également en compte les formes du parfait, la situation se révèle être plus compliquée. Comme le montrent les exemples 6.11 à 6.14 indiquées plus haut la position dans laquelle se placent les MPM dans les formes du parfait correspond, du moins à première vue, à la troisième voire quatrième position, c'est-à-dire après la marque *-û* du parfait ou *-bû* du plus-que-parfait. Les données semblent donc contredire la règle de la seconde position. Elles peuvent apparaître dans les positions I<sub>4</sub> ou I<sub>5</sub> indiquées dans le tableau 6.13.

#### 6.4.3.2 Proposition d'analyse périphrastique des formes du parfait

La contradiction semble néanmoins se résoudre si l'on analyse les formes du parfait comme des formes périphrastiques. Le fait que les formes comportent, du moins sous

---

49. La formalisation de ce complément ainsi que des formes exprimant le parfait a été publiée dans (Walther, 2012a).

forme de traces diachroniques, des éléments qui s'apparentent aux formes auxiliaires du verbe BÛN « être », peut à première vue faire paraître cette hypothèse comme raisonnable. Compte tenu de ce que nous avons indiqué plus haut sur le pouvoir factorisant plus important de la syntaxe par rapport à la morphologie, cela semblerait également être une solution idéale pour obtenir à terme une description globalement plus compacte.

Les désinences D3 employées au parfait sont en effet équivalentes aux formes enclitiques du verbe BÛN « être ». On peut ainsi analyser les formes du parfait comme la combinaison du participe passé (marqué par le RAD2 et la marque du participe *-û*) et les marques du présent (enclitiques) du verbe BÛN « être ».

Dans les formes du plus-que parfait, il est par ailleurs possible de reconnaître les formes du passé du verbe BÛN « être ». Elles sont constituées du RAD2 du BÛN, à savoir *bû*, et des désinences D2 habituelles du passé. Ces formes du passé de BÛN sont combinées au RAD2 des verbes au plus-que-parfait.

Ainsi, les formes du parfait peuvent-elles être vues comme des formes périphrastiques contruites à partir de l'auxiliaire BÛN « être ». Ces formes partagent le même aspect parfait et ne diffèrent que par le temps exprimé, tout comme les formes de l'auxiliaire ne diffèrent qu'en temps (présent *vs.* imparfait). L'hypothèse semble ainsi assez plausible.

Dans une telle approche périphrastique, le placement des MPM semble alors de nouveau analysable en termes de seconde position : elles se placeraient ainsi en seconde position dans l'auxiliaire pour les formes du plus-que-parfait et en seconde position externe, c'est-à-dire comme enclitiques rattachés au participe, pour les formes du parfait. Nous illustrons cette hypothèse par les exemples 6.24 et 6.25. Les crochets représentent les frontières de mot entre verbe lexical et auxiliaire.

6.24 [v xward] [AUX bû=man=in]  
 [manger.RAD2] [AUX.RAD =1PL=3PL]  
 « *Nous les avions mangé(e)s.* » (PST.PRF.TRANS)

6.25 [v xward-û]=man [AUX =in]  
 [manger.RAD2-PTCP]=1PL [=AUX.3PL]  
 « *Nous les avons mangé(e)s.* » (PRS.PRF.TRANS)

### 6.4.3.3 Problèmes de l'analyse périphrastique des formes du parfait du sorani

Cependant, cette analyse pose problème dès lors que l'on prend en compte les faits suivants<sup>50</sup> :

1. Au parfait, les MPM s'attachent effectivement en seconde position externe, c'est-à-dire au premier élément de la construction périphrastique hypothétique, à savoir au participe passé (RAD2 + *û*). En revanche, au plus-que-parfait, elles s'attachent en seconde position interne dans l'auxiliaire. Ainsi les deux constructions, illustrées par les exemples 6.11 et 6.12, semblent correspondre à des constructions dont la nature diffère fondamentalement.

Au lieu de simplifier l'analyse, l'approche périphrastique ne fait ainsi qu'ajouter une complication supplémentaire, à savoir celle de devoir spécifier deux règles différentes pour rendre compte de chacune de ces deux constructions au parfait : une règle de placement externe en seconde position pour le parfait d'une part, et une autre pour le placement interne en seconde position au sein de l'auxiliaire de l'autre.

2. L'analyse périphrastique rend correctement compte de la plupart des formes du parfait. Cependant, elle ne permet pas d'expliquer les ordres inversés entre MPM et marques personnelles (D3 et D2 respectivement) tels qu'ils apparaissent dans les exemples 6.26 et 6.27.

En effet, une opération `compact` supplémentaire serait ici nécessaire pour empêcher la MPM de s'attacher directement au participe passé (en 6.26) ou au radical de l'auxiliaire (en 6.27). Toutefois, si nous adoptons une approche périphrastique de ces formes, une opération morphologique `compact` pose problème en ce cas. Elle aurait en effet à opérer sur des éléments de la périphrase, c'est-à-dire de type mots, type d'opération habituellement réservé à la syntaxe. Ainsi, une telle analyse serait-elle incompatible avec une approche autonome de la morphologie telle que nous l'avions adoptée au chapitre 2.

---

50. Il existe des formes synthétiques pour le subjonctif parfait. Ce fait crée une attente au sens de Brown et al. (2012) à ce qu'il existe également un indicatif parfait et, un cas de périphrase canonique si les formes répondant à l'attente en question ne sont pas de formes synthétiques. Ainsi, l'indicatif du parfait, puisque nous pouvons faire l'hypothèse qu'il pourrait ne pas être synthétique, semble constituer un candidat convaincant d'un cas de périphrase canonique au sens de Brown et al.. En revanche, rien dans le paradigme verbal ne suscite une attente quant à l'existence de formes morphologiquement construites du plus-que-parfait. Le caractère périphrastique du plus-que-parfait serait ainsi au mieux non canonique au sens de Brown et al.. Pourtant, l'on serait également tenté de penser que s'il existe des formes périphrastiques pour le parfait, il est peu probable que les formes du plus-que-parfait puissent être synthétiques. Dans la mesure où, par ailleurs, l'auxiliaire BÛN « être » semble pouvoir se reconnaître aussi bien dans les formes du parfait que dans celles du plus-que-parfait, l'hypothèse du plus-que-parfait périphrastique reste néanmoins probable.

- 6.26 [v xward-û] [AUX =n]=î  
 [v manger.RAD2-PTCP] [AUX =3PL]=3SG  
 « *Il les a mangé(e)s.* » (PRS.PRF.TRANS)
- 6.27 [v xward] [AUX bû-n]=î  
 [v manger.RAD2] [AUX AUX.RAD2 -3PL]=3SG  
 « *Il les avait mangé(e)s.* » (PST.PRF.TRANS.)

3. Une approche lexicaliste (forte) de la morphosyntaxe suppose par ailleurs que la morphologie et la syntaxe soient soumises à des règles dont la nature diffère fondamentalement. Elles appartiennent aussi nécessairement à des composantes différentes de la langue <sup>51</sup>. Par conséquent, il est difficile d'expliquer, dans un modèle lexicaliste fort, quel serait le statut du RAD2 de l'auxiliaire dans la construction du plus-que-parfait, si cette dernière constitue de la périphrase. Les éléments manipulés par la syntaxe ne sont pas soumis aux mêmes règles que les morphes en morphologie. La syntaxe ne peut manipuler que des mots ou des éléments d'une nature similaire à celle des mots, comme les clitiques. Un radical nu, tel que *xward* dans l'exemple 6.27 semble de ce point de vue être un élément syntaxique improbable <sup>52</sup>.
4. De plus, quand les formes du parfait comportent un préfixe négatif, la MPM se retrouve toujours <sup>53</sup> insérée en seconde position interne (c'est-à-dire avant le radical du verbe principal) et non au sein de l'auxiliaire supposé. Ce fait favorise à lui seul très fortement l'analyse synthétique des verbes au parfait, illustrée dans les exemples 6.28 et 6.29.

- 6.28 na=man=xward-û=a  
 NEG=1PL=manger.RAD2-û.PRS.PRF=3SG  
 « *Nous ne l'avons pas mangé(e).* »

51. Cf. chapitre 2.

52. On pourrait néanmoins être tenté de proposer l'existence d'un second type d'infinitifs en sorani qui correspondrait à l'infinitif apocopé du persan. En persan, ces infinitifs correspondent effectivement au radical RAD2 nu d'un point de vue formel (Lazard et al., 2006). Toutefois, cette analyse n'expliquerait pas plus la différence de comportement des MPM au parfait et au plus-que-parfait. En d'autres termes, on s'attendrait toujours à ce que les MPM s'attachent, dans le cas normal, en seconde position externe au parfait comme au plus-que-parfait, c'est-à-dire après le participe et l'infinitif apocopé respectivement.

53. Sauf, évidemment, dans les cas spécifiques liés aux la première et troisième personne du singulier discutés plus haut.



- 6.29 na=**man**=xward-bû=in  
NEG=**1PL**=manger.RAD2-bû.PST.PRF=3SG  
« *Nous ne l'avions pas mangé(e).* »

5. Enfin, on note que le RAD2 du verbe (principal) et l'élément *-bû* marquant le plus-que-parfait ne sont jamais séparés par quelque morphe que ce soit, et ce, quand bien même les MPM peuvent se placer autant immédiatement avant le RAD2 qu'immédiatement après *-bû*. Le fait que ces deux éléments soient à ce point indissociables n'argumente clairement pas en faveur de la présence de deux éléments syntaxiques distincts.

Pour toutes ces raisons, nous pouvons ainsi conclure qu'une analyse synthétique est dans ce cas particulier préférable à une analyse périphrastique. Les MPM du kurde sorani nous donnent ainsi l'exemple d'un cas où une analyse syntaxique n'est pas préférable à une analyse purement morphologique. Dans la sous-section suivante, nous donnons ainsi une proposition d'analyse synthétique du parfait en kurde sorani.

La formalisation des formes verbales du parfait proposée dans (Walther, 2012a) introduit pour rendre compte de ces données une opération *compact* supplémentaire fusionnant systématiquement le radical et les marques *-û* et *-bû* du parfait et du plus-que-parfait. Combinée aux opérations *compact* indiquées plus haut, elle permet de rendre compte du placement des MPM en deuxième position systématique pour l'intégralité des formes du paradigme verbal du sorani.

#### 6.4.4 Conclusion sur la structure à l'interface de la morphologie et de la syntaxe

Le placement des MPM du kurde sorani, à l'interface entre syntaxe et morphologie, constitue un exemple idéal pour étudier la répartition de l'information flexionnelle entre les composantes morphologique et syntaxique de la langue. En effet, les MPM affichent un placement alternant entre positions interne et externe par rapport au verbe. Ils sont ainsi tantôt des enclitiques de seconde position, tantôt des endoclitiques.

Tandis que le statut des MPM (en tant qu'enclitiques ou endoclitiques) est clair dans les phrases comportant des verbes n'exprimant pas le parfait, il est à première vue moins évident dans les phrases avec des formes verbales au parfait. En effet, ces dernières affichent une structure dont l'origine diachronique est clairement périphrastique. Des traces de l'auxiliaire BÛN « *être* » sont aisément reconnaissables. La question d'une analyse

périphrastique ou synthétique en synchronie se pose alors. Si les deux analyses sont possibles, il nous faut par ailleurs trancher en faveur de l'une ou de l'autre.

En vertu du rasoir d'Okham, nous avons dans cette thèse pris le parti de trancher systématiquement en faveur de l'analyse globalement <sup>54</sup> la plus compacte. Or, la syntaxe <sup>55</sup> n'étant, comme la morphonologie, pas soumise aux idiosyncrasies propres à la morphologie, elle constitue *a priori* une composante plus compacte que cette dernière. Ainsi, si une analyse syntaxique comme la périphrase est possible à côté d'une analyse morphologique (synthétique), nous aurions *a priori* tout intérêt à trancher en faveur de l'analyse périphrastique.

Toutefois, cet arbitrage perd tout son sens si l'analyse syntaxique ne contribue pas à alléger la composante morphologique. Or, comme nous l'avons indiqué à la section 6.4.3.3, l'analyse périphrastique suppose un dédoublement de toutes les règles de placement des MPM au parfait : une série de règles pour le placement interne au plus-que-parfait <sup>56</sup> et une série de règles pour le placement externe des formes du parfait. Ainsi, ce dédoublement des règles nous fait perdre l'intégralité du bénéfice que nous aurions pu espérer gagner par le déplacement de l'information flexionnelle dans la composante syntaxique.

Les données du sorani nous montrent ainsi que quand bien même la syntaxe, comme la morphonologie, est *a priori* plus compacte que la morphologie, il existe des cas où l'on ne gagne rien à vouloir alléger la composante morphologique à tout prix. L'équilibre idéal à trouver pour optimiser la compacité réside dans la suppression méticuleuse de toute redondance à travers la description de la langue, toutes composantes confondues.

## 6.5 Composantes de la grammaire et répartition de l'information

Dans les sections 6.2 à 6.4, nous avons montré comment l'information flexionnelle peut être distribuée entre les composantes de la grammaire afin d'obtenir une description la plus compacte possible. Ainsi, dans une approche modulaire intégrée de la grammaire comme celle que nous défendons dans cette thèse, un phénomène dans une langue donnée peut souvent également être modélisé comme étant pris en compte par différentes composantes de la langue.

L'approche que nous proposons dans cette thèse distingue ainsi clairement entre les

---

54. C'est-à-dire toutes composantes de la langue prises en compte.

55. Du moins la syntaxe telle qu'elle est conçue dans un cadre lexicaliste (Chomsky, 1970 ; Bresnan, 1982), cf. la section 2.1.1 du chapitre 2.

56. Nous avons montré que le plus-que-parfait est inanalysable comme une périphrase, du moins dans une approche lexicaliste forte.

composantes de la langue sans pour autant prétendre que la répartition de l'information pour la description d'une langue donnée ne puisse se faire que d'une seule et unique manière. Cependant, nous soutenons qu'il est possible de choisir entre différentes analyses dès lors que l'on adopte des critères de distinction objectifs. Ces critères peuvent être de nature diverse : résultats d'études psycholinguistiques, caractéristiques computationnelles, *etc.* Le critère que nous avons choisi et développé dans ce chapitre est celui de la compacité maximale de l'analyse. Le but recherché est de réduire au maximum la redondance dans la description, et ce de façon intercomposante. Les composantes que nous avons particulièrement prises en compte pour la représentation de la flexion sont celles du lexique, de la morphologie, de la morphonologie et de la phonologie ainsi que de la syntaxe<sup>57</sup>.

La démarche consistant à réduire au maximum la redondance dans la description peut ainsi servir de critère externe et objectif pour guider les choix descriptifs intervenant dans la modélisation constructive (au sens de Blevins (2006, 2012)) d'un (sous-)système flexionnel particulier. Ces choix concernent aussi bien la répartition inter-composante de l'information flexionnelle que les choix internes à la composante morphologique. Ainsi, le fait de poser l'existence de classes flexionnelles ou de catégories flexionnelles supplémentaires plutôt que de décider de l'existence d'un ensemble de flexèmes irréguliers (hétéroclites, décalés, défectifs, *etc.*) aura une influence sur le résultat de l'application de notre métrique de compacité. Le choix pourra ainsi se faire à l'aune de ce critère externe objectif.

Ce n'est qu'une fois que nous aurons ainsi objectivement évalué notre description, qu'il fera sens d'indiquer par des mesures dédiées le taux d'irrégularité du système décrit. Dans notre cas, il s'agira d'appliquer à notre description les mesures de canonicité que nous avons définies au chapitre 4.

Dans le dernier chapitre de cette thèse, nous appliquons donc l'intégralité de nos mesures externe (la métrique de compacité) et internes (les mesures de canonicité) à un sous-système flexionnel particulier, au système verbal du latin.

---

57. Nous ne nous sommes pas plus en détail intéressée aux autres dimensions de la langue comme la sémantique (non lexicale), la pragmatique et les contextes discursifs divers. Dans l'approche que nous adoptons des composantes de la langue, nous considérons que ces composantes s'interfacent avec la syntaxe, mais pas directement avec la morphologie (flexionnelle) qui est l'objet de cette thèse.

## Mesures de canonicité et longueur de description

Dans les chapitres 4 à 6 nous avons présenté le modèle formel de la morphologie flexionnelle  $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$ , son formalisme d'implémentation et les mesures d'évaluations externes permettant de comparer entre elles différentes descriptions morphologiques flexionnelles concurrentes en termes de compacité.  $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$  a été conçu pour servir d'outil de formalisation de la morphologie flexionnelle permettant de rendre compte immédiatement, par la structure de chaque entrée lexicale flexionnelle, des caractéristiques plus ou moins canoniques de paradigmes isolés comme d'un système flexionnel dans son ensemble. Afin de quantifier le degré de canonicité,  $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$  a également été assorti d'une batterie de mesures de canonicité décrites à la section 4.4 du chapitre 4. Elles permettent de situer sur des axes de canonicité gradués les propriétés canoniques des paradigmes flexionnels individuels ou des systèmes flexionnels dans leur ensemble.

Dans ce chapitre, nous comparons deux descriptions concurrentes de la flexion verbale du latin, formalisées en  $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$  et implémentées en  $\text{Alexina}_{\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}}$ . Nous les évaluons en un premier temps en termes de compacité selon la mesure de longueur de description présentée au chapitre 6. Puis nous illustrerons les propriétés non canoniques du système flexionnel verbal latin selon notre analyse en quantifiant l'importance de chaque phénomène non canonique par rapport au système dans son ensemble. Nous utiliserons pour cela les mesures de canonicité dédiées définies au chapitre 4.

## 7.1 Formalisation de la description de la flexion des verbes latins en $\mathcal{PARSL}$

Au chapitre 1, nous avons mis en place une description du système verbal latin reposant sur une analyse dérivationnelle du passif. Dans ce chapitre, nous décrivons la formalisation de cette description en  $\mathcal{PARSL}$ . Cette analyse a fait l'objet d'une implémentation en Alexina $\mathcal{PARSL}$ <sup>1</sup>. Dans la suite de ce chapitre nous appellerons cette analyse « *analyse 3H* ».

Notre analyse repose sur les 2 271 entrées verbales du lexique *Leffa*<sup>2</sup> que nous avons constitué à partir des données contenues dans (Prior & Wohlberg, 1995) complétées par des informations du Wiktionary<sup>3</sup>. Les chiffres de *Leffa* sont indiqués au tableau 1.23 page 85 du chapitre 1. *Leffa* produit à partir des 2 271 entrées intensionnelles 263 677 formes fléchies.

Les tableaux 7.1 à 7.4 pages 411 à 414 présentent huit paradigmes types sur lesquels nous illustrerons tout au long de ce chapitre les propriétés des différentes classes réalisationnelles de nos analyses ainsi que la manifestation de différents phénomènes non canoniques observables, comme les instances de syncrétisme (cases colorées), ou de défektivité et de réalisation périphrastiques (en gris).

### 7.1.1 Définition de la catégorie flexionnelle

L'analyse que nous formalisons portant sur le système verbal du latin, elle repose en un premier temps sur une définition au sein de la grammaire de la catégorie flexionnelle verbale. Compte tenu de l'analyse dérivationnelle que nous avons proposée pour la passivation au chapitre 1, nous pouvons poser une catégorie flexionnelle unique pour tous les verbes. En particulier, nous n'avons plus à distinguer d'un point de vue flexionnel les verbes transitifs passivables des verbes intransitifs non passivables. Dans une analyse flexionnelle du passif latin, ces deux types de verbes auraient en effet eu des propriétés paradigmatiques différentes : les verbes passivables auraient affiché dans leur paradigme des formes actives et passives, les verbes non passivables uniquement des formes actives. Dans notre analyse, la voix ne fait pas partie des traits exprimés par la flexion, de sorte que les

---

1. La description implémentée complète est disponible avec l'inventaire des entrées lexicales traitées sur la page personnelle de l'auteur.

2. *Leffa* est un lexique Alexina $\mathcal{PARSL}$  librement disponible sur la page personnelle de l'auteur.

3. Dictionnaire collaboratif en ligne, semi-structuré, disponible à l'adresse <http://en.wiktionary.org> (valable au 11 juin 2013).



				CONTINEO				CONTINEOR				
IPFV	PRS	1	SG	IND		SBJV		SG	IND		PASS	
				PL	ACT	PL	ACT		PL	IND	PL	SBJV
				<i>continē-o</i>	<i>continē-mus</i>	<i>continē-am</i>	<i>continē-āmus</i>	<i>continē-or</i>	<i>continē-mur</i>	<i>continē-ar</i>	<i>continē-āmur</i>	
		2		<i>continē-s</i>	<i>continē-tis</i>	<i>continē-ās</i>	<i>continē-ātis</i>	<i>continē-ris (-re)</i>	<i>continē-minī</i>	<i>continē-āris (-āre)</i>	<i>continē-āminī</i>	
		3		<i>continē-t</i>	<i>continē-nt</i>	<i>continē-at</i>	<i>continē-ant</i>	<i>continē-tur</i>	<i>continē-ntur</i>	<i>continē-antur</i>	<i>continē-antur</i>	
	PST	1		<i>continē-bam</i>	<i>continē-bāmus</i>	<i>continē-rem</i>	<i>continē-rems</i>	<i>continē-bar</i>	<i>continē-bāmur</i>	<i>continē-bāris (-bāre)</i>	<i>continē-bāminī</i>	
		2		<i>continē-bās</i>	<i>continē-bātis</i>	<i>continē-rēs</i>	<i>continē-rētis</i>	<i>continē-bāris (-bāre)</i>	<i>continē-bāminī</i>	<i>continē-bāris (-bāre)</i>	<i>continē-bāminī</i>	
		3		<i>continē-bat</i>	<i>continē-bant</i>	<i>continē-ret</i>	<i>continē-rent</i>	<i>continē-bātur</i>	<i>continē-bāmur</i>	<i>continē-bāntur</i>	<i>continē-bāmur</i>	
	FUT	1		<i>continē-bo</i>	<i>continē-bimus</i>			<i>continē-bor</i>	<i>continē-bimur</i>	<i>continē-bor</i>	<i>continē-bimur</i>	
		2		<i>continē-bis</i>	<i>continē-bitis</i>			<i>continē-beris (-bere)</i>	<i>continē-binimī</i>	<i>continē-beris (-bere)</i>	<i>continē-binimī</i>	
		3		<i>continē-bit</i>	<i>continē-bant</i>			<i>continē-bitur</i>	<i>continē-bimur</i>	<i>continē-bitur</i>	<i>continē-bimur</i>	
PFV	PRS	1		<i>continū-ī</i>	<i>continū-tinus</i>	<i>continū-erim</i>	<i>continū-erimus</i>	<i>conten̄-tus sum</i>	<i>conten̄-ti sumus</i>	<i>conten̄-tus sim</i>	<i>conten̄-ti simus</i>	
		2		<i>continū-istī</i>	<i>continū-istis</i>	<i>continū-eris</i>	<i>continū-eritis</i>	<i>conten̄-tus es</i>	<i>conten̄-ti estis</i>	<i>conten̄-tus sis</i>	<i>conten̄-ti sitis</i>	
		3		<i>continū-it</i>	<i>continū-erunt (-ere)</i>	<i>continū-erit</i>	<i>continū-erint</i>	<i>conten̄-tus est</i>	<i>conten̄-ti sunt</i>	<i>conten̄-tus sit</i>	<i>conten̄-ti sint</i>	
PST	1			<i>continū-eram</i>	<i>continū-erāmus</i>	<i>continū-issem</i>	<i>continū-issemus</i>	<i>conten̄-tus eram</i>	<i>conten̄-ti erāmus</i>	<i>conten̄-tus essem</i>	<i>conten̄-ti essemus</i>	
		2		<i>continū-erās</i>	<i>continū-erātis</i>	<i>continū-issēs</i>	<i>continū-issetis</i>	<i>conten̄-tus erās</i>	<i>conten̄-ti erātis</i>	<i>conten̄-tus esēs</i>	<i>conten̄-ti essētis</i>	
		3		<i>continū-erat</i>	<i>continū-erant</i>	<i>continū-isset</i>	<i>continū-issent</i>	<i>conten̄-tus erat</i>	<i>conten̄-ti erant</i>	<i>conten̄-tus esset</i>	<i>conten̄-ti essent</i>	
FUT	1			<i>continū-erō</i>	<i>continū-erimus</i>			<i>conten̄-tus erō</i>	<i>conten̄-ti erimus</i>			
		2		<i>continū-eris</i>	<i>continū-eritis</i>			<i>conten̄-tus eris</i>	<i>conten̄-ti eritis</i>			
		3		<i>continū-erit</i>	<i>continū-erint</i>			<i>conten̄-tus erit</i>	<i>conten̄-ti erint</i>			
IMP	PRS	2		<i>continē</i>	<i>continē-te</i>			<i>continē-re</i>	<i>continē-minī</i>			
		2		<i>continē-to</i>	<i>continē-tōte</i>			<i>continē-tor</i>				
		3		<i>continē-to</i>	<i>continē-tōte</i>			<i>continē-tor</i>				
INF	PRS			<i>continē-re</i>	<i>continē-re</i>			<i>continē-ri</i>				
		PST		<i>continū-isse</i>				<i>continē-ntor</i>				
		FUT		<i>conten̄-tūrus/-tura/-tūrum esse</i>				<i>conten̄-tus/-a/-um esse</i>				
PTCP	FUT			<i>continē-ns, -ntis</i>				<i>conten̄-tus/-a/-um</i>				
		PR		<i>conten̄-tūrus/-tura/-tūrum</i>				<i>conten̄-tus/-a/-um</i>				
		FUT		<i>continē-ndū/-nda/-ndum</i>				<i>continē-ndus/-nda/-ndum</i>				
GRV	PST			<i>conten̄-tūrus/-tura/-tūrum</i>				<i>conten̄-tus/-a/-um</i>				
		PR		<i>continē-ndū/-nda/-ndum</i>				<i>continē-ndus/-nda/-ndum</i>				
		FUT		<i>conten̄-tūrus/-tura/-tūrum</i>				<i>conten̄-tus/-a/-um</i>				
GRD	PST			<i>conten̄-tūrus/-tura/-tūrum</i>				<i>conten̄-tus/-a/-um</i>				
		PR		<i>continē-ndū/-nda/-ndum</i>				<i>continē-ndus/-nda/-ndum</i>				
		FUT		<i>conten̄-tūrus/-tura/-tūrum</i>				<i>conten̄-tus/-a/-um</i>				
SUP	PST			<i>conten̄-tūrus/-tura/-tūrum</i>				<i>conten̄-tus/-a/-um</i>				
		PR		<i>continē-ndū/-nda/-ndum</i>				<i>continē-ndus/-nda/-ndum</i>				
		FUT		<i>conten̄-tūrus/-tura/-tūrum</i>				<i>conten̄-tus/-a/-um</i>				

 TABLE 7.2 : Synchrétismes dans les paradigmes des verbes à thème en *e* — exemples de CONTINEO «*contenir*» et CONTINEOR «*être contenu*»

		AUDIO						AUDIOR					
		ACT			PASS			IND			SBJV		
		SG	PL	IND	SG	PL	SG	PL	IND	SG	PL	SG	PL
IPFV	PRS	1	<i>audi-ō</i>	<i>audi-mus</i>	<i>audi-am</i>	<i>audi-āmus</i>	<i>audi-or</i>	<i>audi-mur</i>	<i>audi-ar</i>	<i>audi-ōris (-re)</i>	<i>audi-ōris (-āre)</i>	<i>audi-āris (-āre)</i>	<i>audi-āmur</i>
		2	<i>audi-s</i>	<i>audi-tis</i>	<i>audi-ās</i>	<i>audi-ātis</i>	<i>audi-tur</i>	<i>audi-untur</i>	<i>audi-āris (-re)</i>	<i>audi-ātis (-āre)</i>	<i>audi-ātis (-āre)</i>	<i>audi-āris (-āre)</i>	<i>audi-āmur</i>
		3	<i>audi-t</i>	<i>audi-unt</i>	<i>audi-at</i>	<i>audi-ant</i>	<i>audi-tur</i>	<i>audi-untur</i>	<i>audi-ātis (-āre)</i>	<i>audi-ātis (-āre)</i>	<i>audi-ātis (-āre)</i>	<i>audi-ātis (-āre)</i>	<i>audi-āmur</i>
	PST	1	<i>audi-ēbam</i>	<i>audi-ēbāmus</i>	<i>audi-rēm</i>	<i>audi-rēmus</i>	<i>audi-ēbar</i>	<i>audi-ēbāmur</i>	<i>audi-ēbar</i>	<i>audi-ēbāris (-ēbare)</i>	<i>audi-ēbāris (-ēbare)</i>	<i>audi-ēbāris (-ēbare)</i>	<i>audi-ēbāris (-ēbare)</i>
		2	<i>audi-ēbās</i>	<i>audi-ēbātis</i>	<i>audi-rēs</i>	<i>audi-rētis</i>	<i>audi-ēbātis</i>	<i>audi-ēbāntis</i>	<i>audi-ēbātis</i>	<i>audi-ēbātis (-ēbare)</i>	<i>audi-ēbātis (-ēbare)</i>	<i>audi-ēbātis (-ēbare)</i>	<i>audi-ēbātis (-ēbare)</i>
		3	<i>audi-ēbat</i>	<i>audi-ēbant</i>	<i>audi-ret</i>	<i>audi-rent</i>	<i>audi-ēbātis</i>	<i>audi-ēbantis</i>	<i>audi-ēbātis</i>	<i>audi-ēbātis (-ēbare)</i>	<i>audi-ēbātis (-ēbare)</i>	<i>audi-ēbātis (-ēbare)</i>	<i>audi-ēbātis (-ēbare)</i>
	FUT	1	<i>audi-am</i>	<i>audi-ēmus</i>	<i>audi-erim</i>	<i>audi-erimus</i>	<i>audi-er</i>	<i>audi-erent</i>	<i>audi-ar</i>	<i>audi-eris (-ēre)</i>	<i>audi-eris (-ēre)</i>	<i>audi-eris (-ēre)</i>	<i>audi-eris (-ēre)</i>
		2	<i>audi-ēs</i>	<i>audi-ētis</i>	<i>audi-eris</i>	<i>audi-eritis</i>	<i>audi-eris</i>	<i>audi-eritis</i>	<i>audi-eris</i>	<i>audi-eris (-ēre)</i>	<i>audi-eris (-ēre)</i>	<i>audi-eris (-ēre)</i>	<i>audi-eris (-ēre)</i>
		3	<i>audi-et</i>	<i>audi-ent</i>	<i>audi-erit</i>	<i>audi-erint</i>	<i>audi-eris</i>	<i>audi-erint</i>	<i>audi-erit</i>	<i>audi-eris (-ēre)</i>	<i>audi-eris (-ēre)</i>	<i>audi-eris (-ēre)</i>	<i>audi-eris (-ēre)</i>
PFV	PRS	1	<i>audi-ō</i>	<i>audi-mus</i>	<i>audi-erim</i>	<i>audi-erimus</i>	<i>audi-or</i>	<i>audi-mur</i>	<i>audi-ō</i>	<i>audi-eris</i>	<i>audi-eris</i>	<i>audi-eris</i>	<i>audi-eris</i>
		2	<i>audi-s</i>	<i>audi-tis</i>	<i>audi-eris</i>	<i>audi-eritis</i>	<i>audi-or</i>	<i>audi-untur</i>	<i>audi-eris</i>	<i>audi-eris</i>	<i>audi-eris</i>	<i>audi-eris</i>	<i>audi-eris</i>
		3	<i>audi-t</i>	<i>audi-unt</i>	<i>audi-erit</i>	<i>audi-erint</i>	<i>audi-or</i>	<i>audi-untur</i>	<i>audi-erit</i>	<i>audi-eris</i>	<i>audi-eris</i>	<i>audi-eris</i>	<i>audi-eris</i>
	PST	1	<i>audi-ēbam</i>	<i>audi-ēbāmus</i>	<i>audi-erem</i>	<i>audi-erēmus</i>	<i>audi-erem</i>	<i>audi-erēmur</i>	<i>audi-erem</i>	<i>audi-erem</i>	<i>audi-erem</i>	<i>audi-erem</i>	<i>audi-erem</i>
		2	<i>audi-ēbās</i>	<i>audi-ēbātis</i>	<i>audi-erēs</i>	<i>audi-erētis</i>	<i>audi-erēs</i>	<i>audi-erētis</i>	<i>audi-erēs</i>	<i>audi-erēs</i>	<i>audi-erēs</i>	<i>audi-erēs</i>	<i>audi-erēs</i>
		3	<i>audi-ēbat</i>	<i>audi-ēbant</i>	<i>audi-erāt</i>	<i>audi-erātis</i>	<i>audi-erāt</i>	<i>audi-erātis</i>	<i>audi-erāt</i>	<i>audi-erāt</i>	<i>audi-erāt</i>	<i>audi-erāt</i>	<i>audi-erāt</i>
	FUT	1	<i>audi-erō</i>	<i>audi-erimus</i>	<i>audi-eris</i>	<i>audi-eritis</i>	<i>audi-eris</i>	<i>audi-eritis</i>	<i>audi-eris</i>	<i>audi-eris</i>	<i>audi-eris</i>	<i>audi-eris</i>	<i>audi-eris</i>
		2	<i>audi-eris</i>	<i>audi-eritis</i>	<i>audi-eris</i>	<i>audi-eritis</i>	<i>audi-eris</i>	<i>audi-eritis</i>	<i>audi-eris</i>	<i>audi-eris</i>	<i>audi-eris</i>	<i>audi-eris</i>	<i>audi-eris</i>
		3	<i>audi-erit</i>	<i>audi-erint</i>	<i>audi-erit</i>	<i>audi-erint</i>	<i>audi-erit</i>	<i>audi-erint</i>	<i>audi-erit</i>	<i>audi-eris</i>	<i>audi-eris</i>	<i>audi-eris</i>	<i>audi-eris</i>
IMP	PRS	<i>audi</i>	<i>audi-te</i>	<i>audi-te</i>	<i>audi-te</i>	<i>audi-te</i>	<i>audi-te</i>	<i>audi-te</i>	<i>audi-te</i>	<i>audi-te</i>	<i>audi-te</i>	<i>audi-te</i>	
	FUT	<i>audi-to</i>	<i>audi-tōte</i>	<i>audi-to</i>	<i>audi-tōte</i>	<i>audi-to</i>	<i>audi-tōte</i>	<i>audi-to</i>	<i>audi-to</i>	<i>audi-to</i>	<i>audi-to</i>	<i>audi-to</i>	
		<i>audi-to</i>	<i>audi-tōte</i>	<i>audi-to</i>	<i>audi-tōte</i>	<i>audi-to</i>	<i>audi-tōte</i>	<i>audi-to</i>	<i>audi-to</i>	<i>audi-to</i>	<i>audi-to</i>	<i>audi-to</i>	
INF	PRS	<i>audi-re</i>	<i>audi-eris</i>	<i>audi-re</i>	<i>audi-eris</i>	<i>audi-re</i>	<i>audi-eris</i>	<i>audi-re</i>	<i>audi-re</i>	<i>audi-re</i>	<i>audi-re</i>	<i>audi-re</i>	
	PST	<i>audi-eris</i>	<i>audi-eritis</i>	<i>audi-eris</i>	<i>audi-eritis</i>	<i>audi-eris</i>	<i>audi-eritis</i>	<i>audi-eris</i>	<i>audi-eris</i>	<i>audi-eris</i>	<i>audi-eris</i>	<i>audi-eris</i>	
	FUT	<i>audi-eris</i>	<i>audi-eritis</i>	<i>audi-eris</i>	<i>audi-eritis</i>	<i>audi-eris</i>	<i>audi-eritis</i>	<i>audi-eris</i>	<i>audi-eris</i>	<i>audi-eris</i>	<i>audi-eris</i>	<i>audi-eris</i>	
PTCP	PRS	<i>audi-ens</i>	<i>audi-entis</i>	<i>audi-ens</i>	<i>audi-entis</i>	<i>audi-ens</i>	<i>audi-entis</i>	<i>audi-ens</i>	<i>audi-ens</i>	<i>audi-ens</i>	<i>audi-ens</i>	<i>audi-ens</i>	
	PST	<i>audi-ens</i>	<i>audi-entis</i>	<i>audi-ens</i>	<i>audi-entis</i>	<i>audi-ens</i>	<i>audi-entis</i>	<i>audi-ens</i>	<i>audi-ens</i>	<i>audi-ens</i>	<i>audi-ens</i>	<i>audi-ens</i>	
	FUT	<i>audi-ens</i>	<i>audi-entis</i>	<i>audi-ens</i>	<i>audi-entis</i>	<i>audi-ens</i>	<i>audi-entis</i>	<i>audi-ens</i>	<i>audi-ens</i>	<i>audi-ens</i>	<i>audi-ens</i>	<i>audi-ens</i>	
GRV		<i>audi-ens</i>	<i>audi-entis</i>	<i>audi-ens</i>	<i>audi-entis</i>	<i>audi-ens</i>	<i>audi-entis</i>	<i>audi-ens</i>	<i>audi-ens</i>	<i>audi-ens</i>	<i>audi-ens</i>	<i>audi-ens</i>	
		<i>audi-ens</i>	<i>audi-entis</i>	<i>audi-ens</i>	<i>audi-entis</i>	<i>audi-ens</i>	<i>audi-entis</i>	<i>audi-ens</i>	<i>audi-ens</i>	<i>audi-ens</i>	<i>audi-ens</i>	<i>audi-ens</i>	
		<i>audi-ens</i>	<i>audi-entis</i>	<i>audi-ens</i>	<i>audi-entis</i>	<i>audi-ens</i>	<i>audi-entis</i>	<i>audi-ens</i>	<i>audi-ens</i>	<i>audi-ens</i>	<i>audi-ens</i>	<i>audi-ens</i>	
GRD		<i>audi-ens</i>	<i>audi-entis</i>	<i>audi-ens</i>	<i>audi-entis</i>	<i>audi-ens</i>	<i>audi-entis</i>	<i>audi-ens</i>	<i>audi-ens</i>	<i>audi-ens</i>	<i>audi-ens</i>	<i>audi-ens</i>	
		<i>audi-ens</i>	<i>audi-entis</i>	<i>audi-ens</i>	<i>audi-entis</i>	<i>audi-ens</i>	<i>audi-entis</i>	<i>audi-ens</i>	<i>audi-ens</i>	<i>audi-ens</i>	<i>audi-ens</i>	<i>audi-ens</i>	
		<i>audi-ens</i>	<i>audi-entis</i>	<i>audi-ens</i>	<i>audi-entis</i>	<i>audi-ens</i>	<i>audi-entis</i>	<i>audi-ens</i>	<i>audi-ens</i>	<i>audi-ens</i>	<i>audi-ens</i>	<i>audi-ens</i>	
SUP		<i>audi-ens</i>	<i>audi-entis</i>	<i>audi-ens</i>	<i>audi-entis</i>	<i>audi-ens</i>	<i>audi-entis</i>	<i>audi-ens</i>	<i>audi-ens</i>	<i>audi-ens</i>	<i>audi-ens</i>	<i>audi-ens</i>	
		<i>audi-ens</i>	<i>audi-entis</i>	<i>audi-ens</i>	<i>audi-entis</i>	<i>audi-ens</i>	<i>audi-entis</i>	<i>audi-ens</i>	<i>audi-ens</i>	<i>audi-ens</i>	<i>audi-ens</i>	<i>audi-ens</i>	
		<i>audi-ens</i>	<i>audi-entis</i>	<i>audi-ens</i>	<i>audi-entis</i>	<i>audi-ens</i>	<i>audi-entis</i>	<i>audi-ens</i>	<i>audi-ens</i>	<i>audi-ens</i>	<i>audi-ens</i>	<i>audi-ens</i>	

TABLE 7.3 : Synchrétismes dans les paradigmes des verbes à thème en *i* — exemples d’AUDIO « entendre » et AUDIOR « être entendu »



				DUCCO				DUCCOR			
				ACT		PASS					
				SG	PL	SG	PL	SG	PL	SG	PL
IPFV	PRS	1	SG	<i>duc-ō</i>	IND	<i>duc-am</i>	PL	<i>duc-or</i>	IND	<i>duc-ar</i>	PL
		2		<i>duc-i-s</i>	PL	<i>duc-ās</i>		<i>duc-ri-s (-re)</i>		<i>duc-āris (-āre)</i>	
		3		<i>duc-i-t</i>		<i>duc-at</i>		<i>duc-tur</i>		<i>duc-ātur</i>	
PST	PRS	1		<i>duc-ēbam</i>	PL	<i>duc-rem</i>	PL	<i>duc-ēbar</i>	IND	<i>duc-rem</i>	PL
		2		<i>duc-ēbās</i>		<i>duc-rēs</i>		<i>duc-ēbāris (-ēbare)</i>		<i>duc-rēris (-rēre)</i>	
		3		<i>duc-ēbat</i>		<i>duc-ret</i>		<i>duc-ēbātur</i>		<i>duc-rētur</i>	
FUT	FUT	1		<i>duc-am</i>	PL	<i>duc-erim</i>	PL	<i>duc-ar</i>	IND	<i>duc-eris</i>	PL
		2		<i>duc-ēs</i>		<i>duc-eris</i>		<i>duc-eris (-ere)</i>		<i>duc-eris (-ere)</i>	
		3		<i>duc-et</i>		<i>duc-erit</i>		<i>duc-erit</i>		<i>duc-erit</i>	
PFV	PRS	1		<i>duc-ī</i>	PL	<i>duc-erimus</i>	PL	<i>duc-us sum</i>	IND	<i>duc-us sim</i>	PL
		2		<i>duc-istī</i>		<i>duc-eritis</i>		<i>duc-us es</i>		<i>duc-us sis</i>	
		3		<i>duc-it</i>		<i>duc-erit</i>		<i>duc-us est</i>		<i>duc-us sit</i>	
PST	PST	1		<i>duc-eram</i>	PL	<i>duc-issent</i>	PL	<i>duc-t-us eram</i>	IND	<i>duc-t-us essem</i>	PL
		2		<i>duc-erās</i>		<i>duc-issent</i>		<i>duc-t-us erās</i>		<i>duc-t-us essem</i>	
		3		<i>duc-erat</i>		<i>duc-issent</i>		<i>duc-t-us erat</i>		<i>duc-t-us esset</i>	
FUT	FUT	1		<i>duc-erō</i>	PL	<i>duc-issent</i>	PL	<i>duc-t-us erō</i>	IND	<i>duc-t-us esset</i>	PL
		2		<i>duc-eris</i>		<i>duc-issent</i>		<i>duc-t-us eris</i>		<i>duc-t-us esset</i>	
		3		<i>duc-erit</i>		<i>duc-issent</i>		<i>duc-t-us erit</i>		<i>duc-t-us esset</i>	
IMP	PRS	2		<i>duc</i>	PL	<i>duc-erint</i>	PL	<i>duc-re</i>	IND	<i>duc-re</i>	PL
		2		<i>duc-tō</i>		<i>duc-tōte</i>		<i>duc-tor</i>		<i>duc-tor</i>	
		3		<i>duc-tō</i>		<i>duc-tōte</i>		<i>duc-tor</i>		<i>duc-tor</i>	
INF	PRS			<i>duc-re</i>		<i>duc-re</i>		<i>duc-re</i>		<i>duc-re</i>	
				<i>duc-tō</i>		<i>duc-tōte</i>		<i>duc-tor</i>		<i>duc-tor</i>	
				<i>duc-tō</i>		<i>duc-tōte</i>		<i>duc-tor</i>		<i>duc-tor</i>	
PTCP	FUT			<i>duc-ens</i>		<i>duc-entis</i>		<i>duc-ens</i>		<i>duc-entis</i>	
				<i>duc-ens</i>		<i>duc-entis</i>		<i>duc-ens</i>		<i>duc-entis</i>	
				<i>duc-ens</i>		<i>duc-entis</i>		<i>duc-ens</i>		<i>duc-entis</i>	
GRV	PST			<i>duc-ens</i>		<i>duc-entis</i>		<i>duc-ens</i>		<i>duc-entis</i>	
				<i>duc-ens</i>		<i>duc-entis</i>		<i>duc-ens</i>		<i>duc-entis</i>	
				<i>duc-ens</i>		<i>duc-entis</i>		<i>duc-ens</i>		<i>duc-entis</i>	
GRD	PST			<i>duc-ens</i>		<i>duc-entis</i>		<i>duc-ens</i>		<i>duc-entis</i>	
				<i>duc-ens</i>		<i>duc-entis</i>		<i>duc-ens</i>		<i>duc-entis</i>	
				<i>duc-ens</i>		<i>duc-entis</i>		<i>duc-ens</i>		<i>duc-entis</i>	
SUP	FUT			<i>duc-ens</i>		<i>duc-entis</i>		<i>duc-ens</i>		<i>duc-entis</i>	
				<i>duc-ens</i>		<i>duc-entis</i>		<i>duc-ens</i>		<i>duc-entis</i>	
				<i>duc-ens</i>		<i>duc-entis</i>		<i>duc-ens</i>		<i>duc-entis</i>	

 TABLE 7.4 : Synchrénismes dans les paradigmes des verbes athématisques — exemples de DUCCO « *conduire, mener* » et DUCCOR « *être conduit, mené* »

		ACCIPIO						ACCIPIOR								
		ACT			SBJV			IND			PASS			SBJV		
IPFV		SG	PL	IND	SG	PL	PL	SG	PL	IND	SG	PL	SG	PL	PL	
	PRS	<i>accipi-ō</i>	<i>accipi-mus</i>	<i>accipi-am</i>	<i>accipi-āmus</i>	<i>accipi-erim</i>	<i>accipi-erimus</i>	<i>accipi-or</i>	<i>accipi-mur</i>	<i>accipi-eris (-ere)</i>	<i>accipi-aris (-āre)</i>	<i>accipi-āris (-āre)</i>	<i>accipi-ar</i>	<i>accipi-āris (-āre)</i>	<i>accipi-āmur</i>	
		<i>accipi-s</i>	<i>accipi-tis</i>	<i>accipi-ās</i>	<i>accipi-ātis</i>	<i>accipi-eris</i>	<i>accipi-eritis</i>	<i>accip-eris (-ere)</i>	<i>accipi-mini</i>	<i>accipi-eris (-ere)</i>	<i>accipi-āris (-āre)</i>	<i>accipi-āris (-āre)</i>	<i>accipi-aris (-āre)</i>	<i>accipi-āris (-āre)</i>	<i>accipi-āmini</i>	
	PST	<i>accipi-t</i>	<i>accipi-unt</i>	<i>accipi-at</i>	<i>accipi-ant</i>	<i>accipi-erit</i>	<i>accipi-erint</i>	<i>accipi-tur</i>	<i>accipi-untur</i>	<i>accipi-tur</i>	<i>accipi-tur</i>	<i>accipi-antur</i>	<i>accipi-antur</i>	<i>accipi-antur</i>	<i>accipi-antur</i>	
		<i>accipi-ēbam</i>	<i>accipi-ebāmus</i>	<i>accipi-erem</i>	<i>accipi-rēmus</i>	<i>accipi-erem</i>	<i>accipi-rēmus</i>	<i>accipi-ēbar</i>	<i>accipi-ebāmur</i>	<i>accipi-ēbar</i>	<i>accipi-ēbaris (-āre)</i>	<i>accipi-ēbaris (-āre)</i>	<i>accipi-erem</i>	<i>accipi-erem</i>	<i>accipi-rēmur</i>	
		<i>accipi-ēbās</i>	<i>accipi-ebātis</i>	<i>accipi-rēs</i>	<i>accipi-rētis</i>	<i>accipi-erēs</i>	<i>accipi-erētis</i>	<i>accipi-ēbāt</i>	<i>accipi-ebāmini</i>	<i>accipi-ēbāt</i>	<i>accipi-ēbātis (-āre)</i>	<i>accipi-ēbātis (-āre)</i>	<i>accipi-rēris (-rēre)</i>	<i>accipi-rēris (-rēre)</i>	<i>accipi-rēmini</i>	
	FUT	<i>accipi-ēbat</i>	<i>accipi-ebānt</i>	<i>accipi-ret</i>	<i>accipi-rent</i>	<i>accipi-erit</i>	<i>accipi-erint</i>	<i>accipi-erit</i>	<i>accipi-erint</i>	<i>accipi-erit</i>	<i>accipi-erit</i>	<i>accipi-erint</i>	<i>accipi-erit</i>	<i>accipi-erint</i>	<i>accipi-rentur</i>	
		<i>accipi-am</i>	<i>accipi-ēmus</i>	<i>accipi-erim</i>	<i>accipi-erimus</i>	<i>accipi-erim</i>	<i>accipi-erimus</i>	<i>accipi-am</i>	<i>accipi-ēmur</i>	<i>accipi-am</i>	<i>accipi-erim</i>	<i>accipi-erimus</i>	<i>accipi-erim</i>	<i>accipi-erimus</i>	<i>accipi-erimur</i>	
		<i>accipi-ēs</i>	<i>accipi-ētis</i>	<i>accipi-eris</i>	<i>accipi-eritis</i>	<i>accipi-eris</i>	<i>accipi-eritis</i>	<i>accipi-eris (-ere)</i>	<i>accipi-eris (-ere)</i>	<i>accipi-eris (-ere)</i>	<i>accipi-eris (-ere)</i>	<i>accipi-eris (-ere)</i>	<i>accipi-eris (-ere)</i>	<i>accipi-eris (-ere)</i>	<i>accipi-eris (-ere)</i>	
		<i>accipi-et</i>	<i>accipi-ent</i>	<i>accipi-erit</i>	<i>accipi-erint</i>	<i>accipi-erit</i>	<i>accipi-erint</i>	<i>accipi-erit</i>	<i>accipi-erint</i>	<i>accipi-erit</i>	<i>accipi-erit</i>	<i>accipi-erint</i>	<i>accipi-erit</i>	<i>accipi-erint</i>	<i>accipi-erentur</i>	
PFV	PRS	<i>accip-ī</i>	<i>accip-imus</i>	<i>accip-erim</i>	<i>accip-erimus</i>	<i>accip-erim</i>	<i>accip-erimus</i>	<i>accip-erim</i>	<i>accip-erimus</i>	<i>accip-erim</i>	<i>accip-erimus</i>	<i>accip-erimus</i>	<i>accip-erim</i>	<i>accip-erimus</i>	<i>accip-erimus</i>	
		<i>accip-istī</i>	<i>accip-istis</i>	<i>accip-eris</i>	<i>accip-eritis</i>	<i>accip-eris</i>	<i>accip-eritis</i>	<i>accip-eris</i>	<i>accip-eris</i>	<i>accip-eris</i>	<i>accip-eris</i>	<i>accip-eris</i>	<i>accip-eris</i>	<i>accip-eris</i>	<i>accip-eris</i>	
		<i>accip-it</i>	<i>accip-erunt (-ēre)</i>	<i>accip-erit</i>	<i>accip-erint</i>	<i>accip-erit</i>	<i>accip-erint</i>	<i>accip-erit</i>	<i>accip-erint</i>	<i>accip-erit</i>	<i>accip-erit</i>	<i>accip-erint</i>	<i>accip-erit</i>	<i>accip-erint</i>	<i>accip-erint</i>	
	PST	<i>accip-eram</i>	<i>accip-erāmus</i>	<i>accip-erem</i>	<i>accip-erēmus</i>	<i>accip-erem</i>	<i>accip-erēmus</i>	<i>accip-eram</i>	<i>accip-erāmur</i>	<i>accip-eram</i>	<i>accip-erāmur</i>	<i>accip-erāmur</i>	<i>accip-eram</i>	<i>accip-erāmur</i>	<i>accip-erāmur</i>	
		<i>accip-erās</i>	<i>accip-erātis</i>	<i>accip-erēs</i>	<i>accip-erētis</i>	<i>accip-erēs</i>	<i>accip-erētis</i>	<i>accip-erās</i>	<i>accip-erātis</i>	<i>accip-erās</i>	<i>accip-erās</i>	<i>accip-erās</i>	<i>accip-erās</i>	<i>accip-erās</i>	<i>accip-erās</i>	
		<i>accip-erat</i>	<i>accip-erant</i>	<i>accip-isset</i>	<i>accip-issent</i>	<i>accip-isset</i>	<i>accip-issent</i>	<i>accip-erat</i>	<i>accip-erant</i>	<i>accip-erat</i>	<i>accip-erant</i>	<i>accip-erant</i>	<i>accip-erat</i>	<i>accip-erant</i>	<i>accip-erant</i>	
	FUT	<i>accip-erō</i>	<i>accip-erimus</i>	<i>accip-erim</i>	<i>accip-erimus</i>	<i>accip-erim</i>	<i>accip-erimus</i>	<i>accip-erō</i>	<i>accip-erimus</i>	<i>accip-erō</i>	<i>accip-erimus</i>	<i>accip-erimus</i>	<i>accip-erō</i>	<i>accip-erimus</i>	<i>accip-erimus</i>	
		<i>accip-eris</i>	<i>accip-eritis</i>	<i>accip-eris</i>	<i>accip-eritis</i>	<i>accip-eris</i>	<i>accip-eritis</i>	<i>accip-eris</i>	<i>accip-eris</i>	<i>accip-eris</i>	<i>accip-eris</i>	<i>accip-eris</i>	<i>accip-eris</i>	<i>accip-eris</i>	<i>accip-eris</i>	
		<i>accip-erit</i>	<i>accip-erint</i>	<i>accip-erit</i>	<i>accip-erint</i>	<i>accip-erit</i>	<i>accip-erint</i>	<i>accip-erit</i>	<i>accip-erint</i>	<i>accip-erit</i>	<i>accip-erit</i>	<i>accip-erint</i>	<i>accip-erit</i>	<i>accip-erint</i>	<i>accip-erint</i>	
IMP	PRS	<i>accipe</i>	<i>accipi-te</i>	<i>accipe</i>	<i>accipi-te</i>	<i>accipe</i>	<i>accipi-te</i>	<i>accipe</i>	<i>accipi-te</i>	<i>accipe</i>	<i>accipi-te</i>	<i>accipi-te</i>	<i>accipe</i>	<i>accipi-te</i>	<i>accipi-te</i>	
	FUT	<i>accipi-to</i>	<i>accipi-tōte</i>	<i>accipi-to</i>	<i>accipi-tōte</i>	<i>accipi-to</i>	<i>accipi-tōte</i>	<i>accipi-to</i>	<i>accipi-tōte</i>	<i>accipi-to</i>	<i>accipi-tōte</i>	<i>accipi-tōte</i>	<i>accipi-to</i>	<i>accipi-tōte</i>	<i>accipi-tōte</i>	
		<i>accipi-to</i>	<i>accipi-tōte</i>	<i>accipi-to</i>	<i>accipi-tōte</i>	<i>accipi-to</i>	<i>accipi-tōte</i>	<i>accipi-to</i>	<i>accipi-tōte</i>	<i>accipi-to</i>	<i>accipi-tōte</i>	<i>accipi-tōte</i>	<i>accipi-to</i>	<i>accipi-tōte</i>	<i>accipi-tōte</i>	
INF	PRS	<i>accipe-re</i>	<i>accipi-re</i>	<i>accipe-re</i>	<i>accipi-re</i>	<i>accipe-re</i>	<i>accipi-re</i>	<i>accipe-re</i>	<i>accipi-re</i>	<i>accipe-re</i>	<i>accipi-re</i>	<i>accipi-re</i>	<i>accipe-re</i>	<i>accipi-re</i>	<i>accipi-re</i>	
	PST	<i>accip-isse</i>	<i>accipi-isse</i>	<i>accip-isse</i>	<i>accipi-isse</i>	<i>accip-isse</i>	<i>accipi-isse</i>	<i>accip-isse</i>	<i>accipi-isse</i>	<i>accip-isse</i>	<i>accipi-isse</i>	<i>accipi-isse</i>	<i>accip-isse</i>	<i>accipi-isse</i>	<i>accipi-isse</i>	
	FUT	<i>accip-urus/-ūra/-ūrum esse</i>	<i>accipi-urus/-ūra/-ūrum esse</i>	<i>accip-urus/-ūra/-ūrum esse</i>	<i>accipi-urus/-ūra/-ūrum esse</i>	<i>accip-urus/-ūra/-ūrum esse</i>	<i>accipi-urus/-ūra/-ūrum esse</i>	<i>accip-urus/-ūra/-ūrum esse</i>	<i>accipi-urus/-ūra/-ūrum esse</i>	<i>accip-urus/-ūra/-ūrum esse</i>	<i>accipi-urus/-ūra/-ūrum esse</i>	<i>accipi-urus/-ūra/-ūrum esse</i>	<i>accip-urus/-ūra/-ūrum esse</i>	<i>accipi-urus/-ūra/-ūrum esse</i>	<i>accipi-urus/-ūra/-ūrum esse</i>	
PTCP	PRS	<i>accipi-ēns, -ēntis</i>	<i>accipi-ēns, -ēntis</i>	<i>accipi-ēns, -ēntis</i>	<i>accipi-ēns, -ēntis</i>	<i>accipi-ēns, -ēntis</i>	<i>accipi-ēns, -ēntis</i>	<i>accipi-ēns, -ēntis</i>	<i>accipi-ēns, -ēntis</i>	<i>accipi-ēns, -ēntis</i>	<i>accipi-ēns, -ēntis</i>	<i>accipi-ēns, -ēntis</i>	<i>accipi-ēns, -ēntis</i>	<i>accipi-ēns, -ēntis</i>	<i>accipi-ēns, -ēntis</i>	
	PST	<i>accip-urus/-ūra/-ūrum</i>	<i>accipi-urus/-ūra/-ūrum</i>	<i>accip-urus/-ūra/-ūrum</i>	<i>accipi-urus/-ūra/-ūrum</i>	<i>accip-urus/-ūra/-ūrum</i>	<i>accipi-urus/-ūra/-ūrum</i>	<i>accip-urus/-ūra/-ūrum</i>	<i>accipi-urus/-ūra/-ūrum</i>	<i>accip-urus/-ūra/-ūrum</i>	<i>accipi-urus/-ūra/-ūrum</i>	<i>accipi-urus/-ūra/-ūrum</i>	<i>accip-urus/-ūra/-ūrum</i>	<i>accipi-urus/-ūra/-ūrum</i>	<i>accipi-urus/-ūra/-ūrum</i>	
	FUT	<i>accipi-endus/-enda/-endum</i>	<i>accipi-endus/-enda/-endum</i>	<i>accipi-endus/-enda/-endum</i>	<i>accipi-endus/-enda/-endum</i>	<i>accipi-endus/-enda/-endum</i>	<i>accipi-endus/-enda/-endum</i>	<i>accipi-endus/-enda/-endum</i>	<i>accipi-endus/-enda/-endum</i>	<i>accipi-endus/-enda/-endum</i>	<i>accipi-endus/-enda/-endum</i>	<i>accipi-endus/-enda/-endum</i>	<i>accipi-endus/-enda/-endum</i>	<i>accipi-endus/-enda/-endum</i>	<i>accipi-endus/-enda/-endum</i>	
GRV		<i>accip-um/-ū</i>	<i>accipi-um/-ū</i>	<i>accip-um/-ū</i>	<i>accipi-um/-ū</i>	<i>accip-um/-ū</i>	<i>accipi-um/-ū</i>	<i>accip-um/-ū</i>	<i>accipi-um/-ū</i>	<i>accip-um/-ū</i>	<i>accipi-um/-ū</i>	<i>accipi-um/-ū</i>	<i>accip-um/-ū</i>	<i>accipi-um/-ū</i>	<i>accipi-um/-ū</i>	
GRD		<i>accip-um/-ū</i>	<i>accipi-um/-ū</i>	<i>accip-um/-ū</i>	<i>accipi-um/-ū</i>	<i>accip-um/-ū</i>	<i>accipi-um/-ū</i>	<i>accip-um/-ū</i>	<i>accipi-um/-ū</i>	<i>accip-um/-ū</i>	<i>accipi-um/-ū</i>	<i>accipi-um/-ū</i>	<i>accip-um/-ū</i>	<i>accipi-um/-ū</i>	<i>accipi-um/-ū</i>	
SUP		<i>accip-um/-ū</i>	<i>accipi-um/-ū</i>	<i>accip-um/-ū</i>	<i>accipi-um/-ū</i>	<i>accip-um/-ū</i>	<i>accipi-um/-ū</i>	<i>accip-um/-ū</i>	<i>accipi-um/-ū</i>	<i>accip-um/-ū</i>	<i>accipi-um/-ū</i>	<i>accipi-um/-ū</i>	<i>accip-um/-ū</i>	<i>accipi-um/-ū</i>	<i>accipi-um/-ū</i>	

TABLE 7.5 : Syncrétismes dans les paradigmes des verbes à thème mixte — exemples d'ACCIPIO « accepter » et ACCIPIOR « être accepté »

verbes m-actifs et m-passifs se fléchissent pour le même nombre de traits<sup>4</sup>. Nous n'avons donc pas d'ambiguïté sur l'inventaire de paquets de traits morphosyntaxiques à définir. L'inventaire de ces traits a été présenté au tableau 1.24 du chapitre 1.

Dans notre analyse, une catégorie flexionnelle unique suffit ainsi à rendre compte de toutes les entrées lexicales de notre lexique flexionnel. Les propriétés de cette catégorie flexionnelle sont illustrées par la figure 7.1. La figure 7.1 indique notamment l'inventaire de paquets de traits morphosyntaxiques (PTM) exprimés par la flexion verbale<sup>5</sup>. Un verbe canonique dans notre analyse comporte ainsi 196 PTM, dont 60 formes finies et 6 formes impératives, trois infinitifs,  $3 \times 30$  formes participiales, 30 formes pour l'adjectif verbal, cinq formes pour le gérondif et deux pour le supin. La catégorie flexionnelle verbale se caractérise par ailleurs par un schème flexionnel à quatre niveaux : trois niveaux réalisationnels proprement verbaux<sup>6</sup> (les niveaux 1 à 3 en bleu) auxquels s'ajoute un quatrième niveau permettant la réalisation des marques casuelles des formes participiales (en rouge). Le niveau 4 emprunte des classes flexionnelles appartenant canoniquement à la catégorie flexionnelle adjectivale. Ces classes sont la classe  $Z_{adj-oa}^{exp}$  des adjectifs à flexion dite en *-o* et en *-a* pour les participes futurs et passés et la classe  $Z_{adj-c}^{exp}$  des adjectifs à flexion dite consonantique pour les participes présents. Pour les trois premiers niveaux réalisationnels, les membres canoniques de la catégorie flexionnelle verbale se fléchissent selon une des classes de radicaux référencées au tableau 7.8, un thème et une des classes d'exponence (CLASSE A, CLASSE B, CLASSE C et CLASSE D) définies dans l'analyse du chapitre 1.

## 7.1.2 Définition des trois niveaux réalisationnels principaux

L'analyse 3H stipule, à chaque niveau réalisationnel, un certain nombre de classes réalisationnelles. Ces classes sont au nombre de 17 pour le niveau radical (niveau 1), de quatre pour le niveau thématique (niveau 2) et de quatre pour le niveau d'exponence standard (niveau 3).

4. Par manque, à ce stade, d'un mécanisme de formalisation adapté, traitons pour le moment le cas des formes perfectives périphrastiques comme des cas de défektivité. Cette solution nous semble être moins inapproprié qu'un traitement de la périphrase comme un cas de déficience (définie au chapitre 4, page 222), mais elle reste néanmoins insatisfaisante. Le même traitement est donné aux quelques formes non finies absentes des paradigmes des verbes à flexion m-active (les formes correspondant au PTCP.PST et au GRD), respectivement m-passives (PTCP.PRS, PTCP.FUT, GRV et SUP).

5. La notation  $(1 \vee 2 \vee 3).(SG \vee PL).(PRS \vee PST \vee FUT).(IPFV \vee PFV).(IND \vee SBJV)$  indique ici que MFS comporte toutes les combinaisons possibles des trois personnes (1, 2 et 3) avec les deux nombres (SG et PL), les trois temps (PRS, PST et FUT), les deux aspects (IPFV et PFV) et les deux modes des formes finies (IND et SBJV). Des exemples de ces combinaisons seront ainsi les PTM 1SG.PRS.IPFV.IND ou 3PL.FUT.PFV.SBJV.

6. D'où le nom 3H de cette analyse, le H faisant référence à la description du chapitre 1 qui traite les verbes déponents comme des Hétéroclites.

VERBE	
I-CAT	<i>verbe</i>
MSF	$\left\{ \begin{array}{l} (1\vee2\vee3).(SG\vee PL).(PRS\vee PST).(IPFV\vee PFV).(IND\vee SBJV), \\ (1\vee2\vee3).(SG\vee PL).FUT.(IPFV\vee PFV).IND, \\ 2.(SG\vee PL).PRS.IMP, (2\vee3).(SG\vee PL).FUT.IMP, \\ (PRS\vee PST\vee FUT).INF, GRV.(NOM\vee ACC\vee GEN\vee DAT\vee ABL), \\ (PRS\vee PST\vee FUT).PTCP.(NOM\vee ACC\vee GEN\vee DAT\vee ABL).(SG\vee PL).(M\vee F\vee N), \\ GRD.(NOM\vee ACC\vee GEN\vee DAT\vee ABL).(SG\vee PL).(M\vee F\vee N), SUP.(NOM\vee ACC).SG \end{array} \right\}$
I-PAT	$\left  \begin{array}{l} (Z^s, id), (Z^{th}, id), (Z^{exp}, id), (Z_{adj-oa}^{exp}, id) \\ (Z^s, id), (Z^{th}, id), (Z^{exp}, id), (Z_{adj-c}^{exp}, id) \end{array} \right $

FIGURE 7.1 : Catégorie flexionnelle verbale en latin

Dans cette section nous décrivons les caractéristiques des classes réalisationnelles définies à chaque niveau. Nous présentons en un premier temps les niveaux 2 et 3 et terminons par le niveau 1 qui des trois affiche le plus de diversité de classes réalisationnelles.

Nous donnons quatre exemples d'entrées lexicales aux figures 7.2 à 7.5. Toutes ces entrées présentent à chaque fois une base phonologique identique à leur forme de citation, une catégorie flexionnelle I-CAT verbale, un inventaire de paquets de traits morphosyntaxiques MFS standard, et un schème flexionnel à trois niveaux réalisationnels (radical, thème et exponence) complets auxquels s'ajoutent au niveau 4 une ou deux classes d'exponence empruntées aux classes adjectivales pour la réalisation des marques casuelles des participes.

CONTINEO	
I-PHON	contineo
I-CAT	<i>verbe</i>
MSF	{ <i>standard</i> }
S-STEM	( <i>vide</i> )
S-FORM	( <i>vide</i> )
I-PAT	$\left  \begin{array}{l} (Z_{REG-U-T}^s, id), (Z_e^{th}, id), (Z_{v-A}^{exp}, id), (Z_{adj-oa}^{exp}, id) \\ (Z_{REG-U-T}^s, id), (Z_e^{th}, id), (Z_{v-A}^{exp}, id), (Z_{adj-c}^{exp}, id) \end{array} \right $

FIGURE 7.2 : Analyse 3H : entrée lexicale du flexème CONTINEO « *contenir* »

AUDIOR					
I-PHON	audior				
I-CAT	<i>verbe</i>				
MSF	{ <i>standard</i> }				
S-STEM	( <i>vide</i> )				
S-FORM	( <i>vide</i> )				
I-PAT		( $Z_{\text{REG-V-T}}^s, \text{id}$ ),	( $Z_i^{th}, \text{id}$ )	( $Z_{v-D}^{exp}, \text{id}$ )	( $Z_{adj-oa}^{exp}, \text{id}$ )

FIGURE 7.3 : Analyse 3H : entrée lexicale du flexème AUDIOR « être entendu »

ACCIPIO					
I-PHON	accipio				
I-CAT	<i>verbe</i>				
MSF	{ <i>standard</i> }				
S-STEM	( <i>vide</i> )				
S-FORM	( <i>vide</i> )				
I-PAT		( $Z_{\text{REG-EQ-T}}^s, \text{id}$ ),	( $z_{i,1}^{th}, \text{id}$ )	( $Z_{v-C}^{exp}, \text{id}$ )	( $Z_{adj-oa}^{exp}, \text{id}$ )
		( $Z_{\text{REG-EQ-T}}^s, \text{id}$ ),	( $z_{i,1}^{th}, \text{id}$ )	( $Z_{v-C}^{exp}, \text{id}$ )	( $Z_{adj-c}^{exp}, \text{id}$ )
		( $Z_{\text{REG-EQ-T}}^s, \text{id}$ ),	( $z_{c,2}^{th}, \text{id}$ )	( $Z_{v-C}^{exp}, \text{id}$ )	( $Z_{adj-oa}^{exp}, \text{id}$ )

FIGURE 7.4 : Analyse 3H : entrée lexicale du flexème ACCIPIO « accepter »

### 7.1.2.1 Les classes thématiques

Les classes thématiques sont au nombre de quatre. Elles correspondent aux quatre thèmes habituellement associés aux verbes latins :

- la classe thématique en *a*,
- la classe thématique en *e*,
- la classe thématique en *i* et
- la classe athématique.

ACCIPIOR									
I-PHON	accipior								
I-CAT	<i>verbe</i>								
MSF	{ <i>standard</i> }								
S-STEM	( <i>vide</i> )								
S-FORM	( <i>vide</i> )								
I-PAT	<table style="border-collapse: collapse; width: 100%;"> <tr> <td style="border-right: 1px solid black; padding: 2px 5px;"><math>(Z_{\text{REG-EQ-T}}^s, \text{id})</math>,</td> <td style="padding: 2px 5px;"><math>(z_{i,1}^{th}, \text{id})</math></td> <td style="padding: 2px 5px;"><math>(Z_{v-D}^{exp}, \text{id})</math></td> <td style="padding: 2px 5px;"><math>(Z_{adj-oa}^{exp}, \text{id})</math></td> </tr> <tr> <td style="border-right: 1px solid black; padding: 2px 5px;"><math>(Z_{\text{REG-EQ-T}}^s, \text{id})</math>,</td> <td style="padding: 2px 5px;"><math>(z_{c,2}^{th}, \text{id})</math></td> <td style="padding: 2px 5px;"><math>(Z_{v-D}^{exp}, \text{id})</math></td> <td style="padding: 2px 5px;"><math>(Z_{adj-oa}^{exp}, \text{id})</math></td> </tr> </table>	$(Z_{\text{REG-EQ-T}}^s, \text{id})$ ,	$(z_{i,1}^{th}, \text{id})$	$(Z_{v-D}^{exp}, \text{id})$	$(Z_{adj-oa}^{exp}, \text{id})$	$(Z_{\text{REG-EQ-T}}^s, \text{id})$ ,	$(z_{c,2}^{th}, \text{id})$	$(Z_{v-D}^{exp}, \text{id})$	$(Z_{adj-oa}^{exp}, \text{id})$
$(Z_{\text{REG-EQ-T}}^s, \text{id})$ ,	$(z_{i,1}^{th}, \text{id})$	$(Z_{v-D}^{exp}, \text{id})$	$(Z_{adj-oa}^{exp}, \text{id})$						
$(Z_{\text{REG-EQ-T}}^s, \text{id})$ ,	$(z_{c,2}^{th}, \text{id})$	$(Z_{v-D}^{exp}, \text{id})$	$(Z_{adj-oa}^{exp}, \text{id})$						

FIGURE 7.5 : Analyse 3H : entrée lexicale du flexème ACCIPIOR « être accepté »

Dans la formalisation  $\mathbb{P}\mathbb{R}\mathbb{R}\mathbb{S}\mathbb{L}$  que nous donnons de notre analyse du latin, les classes thématiques THÈME-A, THÈME-E, THÈME-I et THÈME-C seront respectivement notées  $Z_a^{th}$ ,  $Z_e^{th}$ ,  $Z_i^{th}$  et  $Z_c^{th}$ . Deux exemples de verbes pour chacune de ces classes sont donnés aux tableaux 7.1 à 7.4.

On constate pour les verbes CONTINEO et AUDIOR des figures 7.2 à 7.3 qu'ils se fléchissent au niveau thématique (niveau 2) selon une classe thématique unique, en l'occurrence les classes  $Z_e^{th}$  et  $Z_i^{th}$ . En revanche, les verbes ACCIPIO et ACCIPIOR illustrés par le tableau 7.5 constituent des cas d'hétéroclise au niveau thématique. Dans la représentation des entrées lexicales aux figures 7.4 et 7.5 associées à chacun de ces flexèmes, cette hétéroclise se reflète dans la présence dans les sous-schémas des schémas des flexèmes en question de deux zones de niveau 2 appartenant à deux classes thématiques distinctes :  $z_{i,1}^{th}$  appartenant à  $Z_i^{th}$  et  $z_{c,2}^{th}$  appartenant à  $Z_c^{th}$ .

Le tableau 7.6 résume les différentes classes thématiques avec le nombre de flexèmes associés à chacune d'entre elles ainsi que le nombre de flexèmes hétéroclites au niveau thématique comme ACCIPIO « *accepter* ».

### 7.1.2.2 Les classes d'exponence

Dans notre analyse du chapitre 1, nous avons défini quatre classes d'exponence verbales :

- Deux classes pour les verbes à désinences m-actives :
  - la CLASSE-A pour les verbes à THÈME-A ou THÈME-E
  - la CLASSE-C pour les verbes à THÈME-I, THÈME-C ou thème mixte.

CLASSE	NOMBRE	SPÉCIFICITÉS	EXEMPLE	TRADUCTION
$Z_a^{th}$	803	thème en <i>-a</i>	AMO	« aimer »
$Z_e^{th}$	244	thème en <i>-e</i>	CONTINEO	« contenir »
$Z_i^{th}$	108	thème en <i>-i</i>	AUDIO	« entendre »
$Z_c^{th}$	926	thème en <i>-c</i>	DUCO	« conduire, mener »
mixtes	126	hétéroclites $Z_i^{th}$ et $Z_c^{th}$	ACCIPIO	« accepter »

TABLE 7.6 : Classes thématiques définies au niveau 2 de l'analyse 3H

- Deux classes pour les désinences m-passives, notamment pour les verbes dérivés des verbes appartenant aux classes A et C.
  - la CLASSE-B pour les verbes à THÈME-A ou THÈME-E
  - la CLASSE-D pour les verbes à THÈME-I, THÈME-C ou thème mixte.

Dans notre analyse nous noterons ces quatre classes  $Z_{v-A}^{exp}$ ,  $Z_{v-B}^{exp}$ ,  $Z_{v-C}^{exp}$  et  $Z_{v-D}^{exp}$ .

Outre la différence entre désinences m-actives et m-passives entre les classes  $Z_{v-A}^{exp}$  et  $Z_{v-C}^{exp}$  d'une part et  $Z_{v-B}^{exp}$  et  $Z_{v-D}^{exp}$  de l'autre, ces quatre classes présentent également les différences suivantes : les marques du FUT.IPFV.IND et du PST.IPFV.IND sont en *-bi-* et *-ba* pour  $Z_{v-A}^{exp}$  et  $Z_{v-B}^{exp}$ , tandis qu'elles sont en *-e*<sup>7</sup> et *-eba* pour  $Z_{v-C}^{exp}$  et  $Z_{v-D}^{exp}$ . Par ailleurs  $Z_{v-A}^{exp}$  et  $Z_{v-B}^{exp}$  forment leur participe présent, leur gérondif et leur adjectif verbal en *-ns/-nd* (PTCP.PRS) et en *-nd* (GRV et GRD) là où  $Z_{v-C}^{exp}$  et  $Z_{v-D}^{exp}$  les forment en *-ens/-end* (PTCP.PRS) et en *-end* (GRV et GRD).

Enfin, les verbes appartenant à  $Z_{v-A}^{exp}$  et  $Z_{v-C}^{exp}$  ne réalisent pas de formes pour le participe passé et l'adjectif verbal et sont périphrastiques à l'infinitif futur. Les verbes appartenant à  $Z_{v-B}^{exp}$  et  $Z_{v-D}^{exp}$  ne comportent pas de règles de réalisation pour les participes présent et futur et au gérondif et sont périphrastiques aux infinitifs passé et futur ainsi que pour l'ensemble des formes finies perfectives. Ces cas de défektivité sont indiqués en gris foncé dans les tableaux 7.1 à 7.4. Les instances de périphrase y sont indiquées en gris clair.

Le tableau 7.7 résume les différentes classes d'exponence dans l'analyse 3H avec le nombre de flexèmes associés à chacune d'entre elles ainsi que le nombre de flexèmes hétéroclites de type  $Z_{v-A}^{exp}$  et  $Z_{v-B}^{exp}$  comme OPINOR « penser » et  $Z_{v-C}^{exp}$  et  $Z_{v-D}^{exp}$  comme LOQUOR « parler ».

7. En *-a-* pour le 1SG.FUT.IPFV.IND.

CLASSE	NOMBRE	SPÉCIFICITÉS	EXEMPLE	TRADUCTION
$Z_{v-A}^{exp}$	504	<ul style="list-style-type: none"> <li>– désinences m-actives,</li> <li>– PRS.IPFV.IND en <i>-ba-</i>,</li> <li>– FUT.IPFV en <i>-bi-</i></li> <li>– GRV en <i>-nd</i></li> <li>– pas de PTCP.PST, GRD</li> <li>– PTCP.PRS en <i>-ns/-nd</i></li> <li>– INF.FUT périphrastique</li> </ul>	AMO	« <i>aimer</i> »
$Z_{v-B}^{exp}$	447	<ul style="list-style-type: none"> <li>– désinences m-actives,</li> <li>– PRS.IPFV.IND en <i>-ba-</i>,</li> <li>– FUT.IPFV en <i>-bi-</i></li> <li>– GRD en <i>-nd</i></li> <li>– pas de PTCP.PRS, PTCP.FUT, GRV</li> <li>– PFV et INF.PST périphrastique</li> </ul>	AMOR	« <i>être aimé</i> »
$Z_{v-C}^{exp}$	561	<ul style="list-style-type: none"> <li>– désinences m-actives,</li> <li>– PRS.IPFV.IND en <i>-eba-</i>,</li> <li>– FUT.IPFV en <i>-e- (-a-)</i></li> <li>– GRV en <i>-end</i></li> <li>– pas de PTCP.PST, GRD</li> <li>– PTCP.PRS en <i>-ens/-end</i></li> <li>– INF.FUT périphrastique</li> </ul>	AUDIO	« <i>entendre</i> »
$Z_{v-D}^{exp}$	529	<ul style="list-style-type: none"> <li>– désinences m-actives,</li> <li>– PRS.IPFV.IND en <i>-eba-</i>,</li> <li>– FUT.IPFV en <i>-e- (-a-)</i></li> <li>– GRD en <i>-end</i></li> <li>– pas de PTCP.PRS, PTCP.FUT, GRV</li> <li>– PFV et INF.PST périphrastique</li> </ul>	DUCOR	« <i>être entendu</i> »
hétéroclites	96	$Z_{v-A}^{exp}$ et $Z_{v-B}^{exp}$	OPINOR	« <i>penser</i> »
hétéroclites	70	$Z_{v-C}^{exp}$ et $Z_{v-D}^{exp}$	LOQUOR	« <i>parler</i> »

TABLE 7.7 : Classes d'exponence définies au niveau 2 de l'analyse 3H



## 7.1.2.3 Les classes de radicaux

Notre description définit également 17 classes de radicaux recensées au tableau 7.8 (avec à chaque fois le nombre de flexèmes associé à chacune d'entre elles). Ces classes se différencient les unes de autres par la manière dont sont formés le RAD2, respectivement le RAD3. Elles utilisent ainsi une zone de radical pour la formation du RAD1 partagée par toutes les classes, sept zones de radicaux pour la formation du RAD2 et sept pour la formation du RAD3. La définition de ces zones s'appuie sur des généralisations sur les procédés morphologiques employés pour construire les radicaux RAD2 et RAD3. Ces généralisations ont été conçues de manière à ce que chaque zone soit employée par un nombre minimum de flexèmes ; dans notre analyse nous avons pour cela placé le seuil à au moins 12 flexèmes associés. Les autres types de construction de radicaux ont été considérés comme relevant de la supplétion. Les radicaux correspondants sont alors répertoriés dans les listes de radicaux supplétifs des flexèmes concernés. Les noms des classes de radicaux au tableau 7.8 représentent à chaque fois la combinaison des zones employées pour les construire. Par exemple  $Z_{\text{REG-V-T}}^s$  est une classe construisant son RAD2 par l'ajout de *-v* au RAD1<sup>8</sup> et son RAD3 par l'ajout de *-t* au RAD1. La classe  $Z_{\text{H-EQ-SS}}^s$  construit son RAD2 par égalité avec le RAD1 et son RAD3 avec la règle stipulant l'ajout de *-ss* au RAD1.

Les espaces partitionnants définissant chacune de ces zones sont ceux définis au chapitre 1, à savoir :

**EP-RAD1** L'espace partitionnant EP-RAD1 couvre tous les paquets de traits morphosyntaxiques (PTM) comportant le trait IPFV pour les formes finies, ainsi que celle de l'infinitif présent (PRS.INF), du participe présent (PRS.PTCP) et du gérondif (GRV).

**EP-RAD2** L'espace partitionnant EP-RAD2 regroupe tous les PTM comportant le trait PFV pour les formes finies et l'infinitif passé (PST.INF).

**EP-RAD3** L'espace partitionnant EP-RAD3 regroupe les PTM de l'infinitif et du participe futurs (FUT.INF et FUT.PTCP) et du supin (SUP).

Aux zones régulières s'ajoutent deux zones vides pour la classe des radicaux entièrement supplétifs au RAD2 et au RAD3, ainsi que des espaces partitionnants spécifiques à la supplétion de radicaux des verbes SUM « être » et EO « aller » ainsi que de leurs dérivés<sup>9</sup>.

8. Avec éventuellement quelques changements morphologiques réguliers en plus. Cela vaut également pour les autres zones définies.

9. POSSUM « pouvoir », OBSUM « être opposé », ADSUM « être présent », PRAESUM « être responsable », OBEO « rencontrer », ADEO « approcher », etc.

10. Ce nombre inclut également les verbes défectifs pour le RAD2 mais dont le RAD3 est construit selon la classe  $Z_{\text{REG-V-T}}^s$ . Il en va de même pour les verbes à RAD2 supplétif (pour lesquels les règles de construction

CLASSE	NOMBRE	SPÉCIFICITÉS	RAD1	RAD2	RAD3	TRADUCTION
$Z_{\text{REG-V-T}}^s$	1 298 <sup>10</sup>	RAD2 en <i>-v</i> , RAD3 en <i>-t</i>	ama-	amav-	amat-	« <i>aimer</i> »
$Z_{\text{REG-EQ-T}}^s$	206	RAD2 = RAD1, RAD3 en <i>-t</i>	ven-	ven-	vent-	« <i>venir</i> »
$Z_{\text{REG-EQ-S}}^s$	160	RAD2 = RAD1, RAD3 en <i>-s</i>	prand-	prand-	prans-	« <i>déjeuner</i> »
$Z_{\text{REG-U-T}}^s$	122	RAD2 en <i>-u</i> , RAD3 en <i>-t</i>	aper-	aperu-	apert-	« <i>ouvrir</i> »
$Z_{\text{REG-X-T}}^s$	92	RAD2 en <i>-x</i> , RAD3 en <i>-t</i>	luc-	lux-	luct-	« <i>briller</i> »
$Z_{\text{REG-S-SS}}^s$	75	RAD2 en <i>-s</i> , RAD3 en <i>-ss</i>	permitt-	permis-	permiss-	« <i>permettre</i> »
$Z_{\text{H-S-T}}^s$	51	RAD2 en <i>-s</i> , RAD3 en <i>-t</i>	nub-	nubs-	nubt-	« <i>épouser</i> »
$Z_{\text{H-S-S}}^s$	45	RAD2 en <i>-s</i> , RAD3 en <i>-s</i>	accend-	accens-	accens-	« <i>embraser</i> »
$Z_{\text{H-EQ-VT}}^s$	20	RAD2 = RAD1, RAD3 en <i>-t</i> avec chute du <i>v</i> final	foc-	fov-	fot-	« <i>adorer</i> »
$Z_{\text{REG-UL}}^s$	20	RAD2 en <i>ul</i> , RAD3 en <i>-uls</i>	impell-	impul-	impuls-	« <i>heurter</i> »
$Z_{\text{REG-X-X}}^s$	18	RAD2 en <i>-x</i> , RAD3 en <i>-x</i>	fig-	fix-	fix-	« <i>planter</i> »
$Z_{\text{H-SS-SS}}^s$	15	RAD2 en <i>-ss</i> , RAD3 en <i>-ss</i>	concut-	concuss-	concuss-	« <i>secouer</i> »
$Z_{\text{H-EQ-EQ}}^s$	9	RAD2 = RAD3 = RAD1	vis-	vis-	vis-	« <i>regarder</i> »
$Z_{\text{H-SS-T}}^s$	6	RAD2 en <i>-ss</i> , RAD3 en <i>-t</i>	ger-	gess-	gest-	« <i>tourner</i> »
$Z_{\text{H-EQ-SS}}^s$	6	RAD2 = RAD1, RAD3 en <i>-ss</i>	pand-	pand-	pass-	« <i>étaler</i> »
$Z_{\text{H-X-EQ}}^s$	3	RAD2 en <i>-x</i> , RAD3 = RAD1	pect-	pex-	pect-	« <i>peigner</i> »
$Z_{\text{SUPPL}}^s$	125	classe vide pour les verbes entièrement supplétifs, dont : – les verbes à espace partitionnant spécifique aux dérivés de EO – les verbes à espace partitionnant spécifique aux dérivés de SUM			e-, i-	« <i>aller</i> »  « <i>être</i> »

TABLE 7.8 : Classes de radicaux définies

### 7.1.3 Combinaison inter-niveaux des classes réalisationnelles

Les classes des différents niveaux réalisationnels de l'analyse 3H se combinent selon les flexèmes concernés comme indiqué à la figure 7.6. Chaque forme (non supplétive) est le résultat de l'application d'une règle réalisationnelle par niveau. Ces règles appartiennent aux classes indiquées dans les boîtes de la figure 7.6 : une des 17 classes de radicaux au niveau 1, une des quatre classes thématiques (THÈME A à THÈME C) au niveau 2 et une des quatre classes d'exponences (CLASSE A à CLASSE D) au niveau 3. Au niveau 4 s'y ajoutent pour les participes, l'adjectif verbal, le supin et le gérondif des règles pour la réalisation du cas et du nombre selon l'une des deux classes d'exponence adjectivales CLASSE ADJ-C et CLASSE ADJ-OA. La figure 7.6 illustre en particulier une association systématique entre l'appartenance à une des classes thématiques  $Z_a^{th}$  ou  $Z_e^{th}$  avec une classe d'exponence  $Z_{v-A}^{exp}$  ou  $Z_{v-B}^{exp}$ , de même qu'entre les classes  $Z_i^{th}$  et  $Z_c^{th}$  avec les classes  $Z_{v-C}^{exp}$  ou  $Z_{v-D}^{exp}$ . Les classes adjectivales du niveau 4 seront quant à elles compatibles avec toutes les classes du niveau 3 pour  $Z_{adj-oa}^{exp}$  et seules les classes  $Z_{v-A}^{exp}$  ou  $Z_{v-C}^{exp}$  pour  $Z_{adj-c}^{exp}$ . Cela relève du fait que  $Z_{adj-c}^{exp}$  sert à la réalisation des participes présents que l'on ne rencontre que pour les flexèmes à désinences m-actives, c'est-à-dire se fléchissant selon les classes  $Z_{v-A}^{exp}$  ou  $Z_{v-C}^{exp}$ .

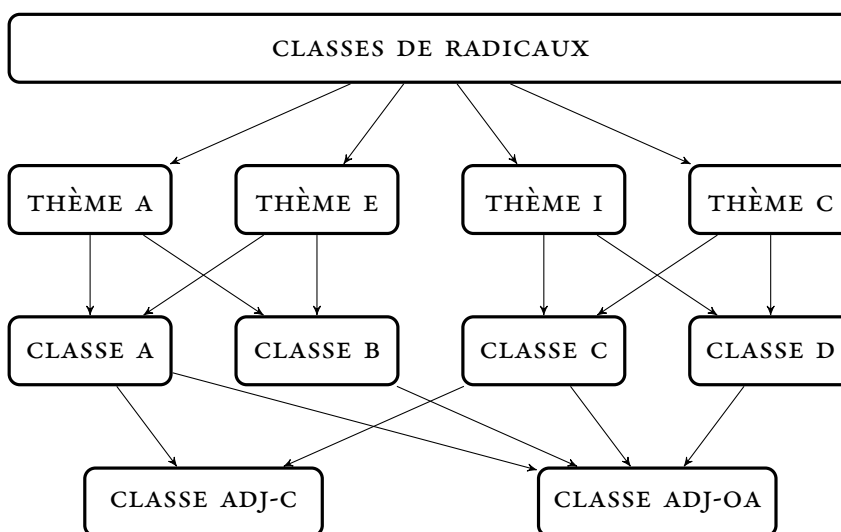


FIGURE 7.6 : Classes réalisationnelles dans l'analyse 3H

du RAD2 sont court-circuitées par la spécification d'un radical dans le lexique). En attribuant des classes existantes à des verbes défectifs ou supplétifs, nous avons cherché à factoriser au maximum l'emploi des classes définies. Le même choix a ainsi été fait pour les classes  $Z_{REG-EQ-S}^s$ ,  $Z_{REG-U-T}^s$ ,  $Z_{REG-S-SS}^s$ ,  $Z_{REG-EQ-T}^s$ ,  $Z_{H-S-S}^s$ ,  $Z_{REG-EQie-T}^s$ ,  $Z_{REG-EQ-VT}^s$ ,  $Z_{REG-X-X}^s$  et  $Z_{H-EQ-EQ}^s$ .

### 7.1.4 Représentation de la dérivation passive

Le point principal de notre analyse du chapitre 1 consiste à donner une interprétation dérivationnelle au passif en latin. Ainsi, certains flexèmes à désinences m-actives peuvent être dérivés pour former les flexèmes correspondants à désinences m-passives. Formellement, cette dérivation se manifeste essentiellement par un changement du schème flexionnel I-PAT de l'entrée au niveau exposant<sup>11</sup>. Nous illustrons ce changement par les entrées lexicales des couples AMO/AMOR « *aimer/être aimé* » aux figures 7.7 et 7.8 d'une part et des couples à thème hétéroclite ACCIPIO/ACCIPIOR « *accepter/être accepté* » aux figures 7.9 et 7.10 de l'autre.

AMO									
I-PHON	amo								
I-CAT	<i>verbe</i>								
MSF	{ <i>standard</i> }								
S-STEM	( <i>vide</i> )								
S-FORM	( <i>vide</i> )								
I-PAT	<table style="border-collapse: collapse; width: 100%;"> <tr> <td style="border-right: 1px solid black; padding: 2px 5px;"><math>(Z_{\text{REG-V-T}}^s, \text{id})</math>,</td> <td style="padding: 2px 5px;"><math>(Z_a^{\text{th}}, \text{id})</math></td> <td style="padding: 2px 5px;"><math>(Z_{v-A}^{\text{exp}}, \text{id})</math></td> <td style="padding: 2px 5px;"><math>(Z_{\text{adj-oa}}^{\text{exp}}, \text{id})</math></td> </tr> <tr> <td style="border-right: 1px solid black; padding: 2px 5px;"><math>(Z_{\text{REG-V-T}}^s, \text{id})</math>,</td> <td style="padding: 2px 5px;"><math>(Z_a^{\text{th}}, \text{id})</math></td> <td style="padding: 2px 5px;"><math>(Z_{v-A}^{\text{exp}}, \text{id})</math></td> <td style="padding: 2px 5px;"><math>(Z_{\text{adj-c}}^{\text{exp}}, \text{id})</math></td> </tr> </table>	$(Z_{\text{REG-V-T}}^s, \text{id})$ ,	$(Z_a^{\text{th}}, \text{id})$	$(Z_{v-A}^{\text{exp}}, \text{id})$	$(Z_{\text{adj-oa}}^{\text{exp}}, \text{id})$	$(Z_{\text{REG-V-T}}^s, \text{id})$ ,	$(Z_a^{\text{th}}, \text{id})$	$(Z_{v-A}^{\text{exp}}, \text{id})$	$(Z_{\text{adj-c}}^{\text{exp}}, \text{id})$
$(Z_{\text{REG-V-T}}^s, \text{id})$ ,	$(Z_a^{\text{th}}, \text{id})$	$(Z_{v-A}^{\text{exp}}, \text{id})$	$(Z_{\text{adj-oa}}^{\text{exp}}, \text{id})$						
$(Z_{\text{REG-V-T}}^s, \text{id})$ ,	$(Z_a^{\text{th}}, \text{id})$	$(Z_{v-A}^{\text{exp}}, \text{id})$	$(Z_{\text{adj-c}}^{\text{exp}}, \text{id})$						

FIGURE 7.7 : Analyse 3H : entrée lexicale du flexème de la base AMO « *aimer* »

Les figures 7.7 à 7.10 montrent ainsi que les entrées lexicales de verbes appartenant de façon standard à la classe  $Z_{v-A}^{\text{exp}}$  acquièrent lors de leur dérivation passive un schème flexionnel comportant la classe  $Z_{v-B}^{\text{exp}}$ . Le correspondant m-passif d'un verbe de la classe  $Z_{v-C}^{\text{exp}}$  reçoit la classe  $Z_{v-D}^{\text{exp}}$ . Ces changements sont indiqués en rouge dans les entrées des figures 7.7 à 7.10. On note que AMO et son dérivé AMOR partagent toujours la même classe thématique  $Z_a^{\text{th}}$  (indiquée en bleu). De la même manière, ACCIPIO et son dérivé ACCIPIOR partagent les mêmes propriétés au niveau thématique : tous deux sont hétéroclites et emploient des zones des classes thématiques  $Z_i^{\text{th}}$  et  $Z_c^{\text{th}}$  (indiquées en vert). Ainsi, au cours de la dérivation passive, l'appartenance thématique, y compris dans ses propriétés non canoniques, se maintient. L'appartenance à la classe de radical reste quant à elle également inchangée. Le seul changement est le passage à une autre classe d'exponence,

11. Nous ne représentons ici que les propriétés flexionnelles des lexèmes base et dérivé. La représentation de l'opération dérivationnelle complète dépasse à ce stade les réalités formalisées par  $\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}$ .

AMOR	
I-PHON	amor
I-CAT	<i>verbe</i>
MSF	{ <i>standard</i> }
S-STEM	( <i>vide</i> )
S-FORM	( <i>vide</i> )
I-PAT	( $Z_{\text{REG-V-T}}^s, \text{id}$ ), ( $Z_a^{th}, \text{id}$ ) ( $Z_{v-B}^{exp}, \text{id}$ ) ( $Z_{adj-oa}^{exp}, \text{id}$ )

FIGURE 7.8 : Analyse 3H : entrée lexicale du flexème AMOR « être aimé » dérivé de AMO « aimer »

ACCIPIO	
I-PHON	accipio
I-CAT	<i>verbe</i>
MSF	{ <i>standard</i> }
S-STEM	( <i>vide</i> )
S-FORM	( <i>vide</i> )
I-PAT	( $Z_{\text{REG-EQ-T}}^s, \text{id}$ ), ( $z_{i,1}^{th}, \text{id}$ ) ( $Z_{v-C}^{exp}, \text{id}$ ) ( $Z_{adj-oa}^{exp}, \text{id}$ ) ( $Z_{\text{REG-EQ-T}}^s, \text{id}$ ), ( $z_{i,1}^{th}, \text{id}$ ) ( $Z_{v-C}^{exp}, \text{id}$ ) ( $Z_{adj-c}^{exp}, \text{id}$ ) ( $Z_{\text{REG-EQ-T}}^s, \text{id}$ ), ( $z_{c,2}^{th}, \text{id}$ ) ( $Z_{v-C}^{exp}, \text{id}$ ) ( $Z_{adj-oa}^{exp}, \text{id}$ )

FIGURE 7.9 : Analyse 3H : entrée lexicale du flexème ACCIPIO « accepter »

les classes étant compatibles deux à deux pour cette relation de dérivation :

- $Z_{v-A}^{exp}$  donne  $Z_{v-B}^{exp}$  lors de la passivation,
- $Z_{v-C}^{exp}$  donne  $Z_{v-D}^{exp}$  lors de la passivation.

Ce parallélisme systématique pose par ailleurs un certain nombre de questions qui sont à l'origine du développement de notre analyse concurrente 4H décrite plus bas.

### 7.1.5 Représentation de verbes déponents

Comme nous l'avons montré au chapitre 1, dans une approche dérivationnelle du passif en latin, les verbes déponents sont, en ce qu'ils possèdent à la fois des participes présent,

ACCIPIOR					
I-PHON	accipior				
I-CAT	<i>verbe</i>				
MSF	{ <i>standard</i> }				
S-STEM	( <i>vide</i> )				
S-FORM	( <i>vide</i> )				
I-PAT		( $Z_{\text{REG-EQ-T}}^s, \text{id}$ ),	( $z_{i,1}^{th}, \text{id}$ )	( $Z_{v-D}^{exp}, \text{id}$ )	( $Z_{adj-oa}^{exp}, \text{id}$ )
		( $Z_{\text{REG-EQ-T}}^s, \text{id}$ ),	( $z_{c,2}^{th}, \text{id}$ )	( $Z_{v-D}^{exp}, \text{id}$ )	( $Z_{adj-oa}^{exp}, \text{id}$ )

FIGURE 7.10 : Analyse 3H : entrée lexicale du flexème ACCIPIOR « être accepté »

futur et passé, des infinitifs présent et passé et le gérondif, l'adjectif verbal et le supin, des verbes présentant des formes surnuméraires par rapport aux autres verbes à désinences m-passives<sup>12</sup>. Puisque leurs paradigmes comportent ainsi des formes qui sont pour la plupart produites par les classes flexionnelles  $Z_{v-B}^{exp}$  ou  $Z_{v-D}^{exp}$  mais pour certaines par les classes  $Z_{v-A}^{exp}$  et  $Z_{v-C}^{exp}$ , les verbes déponents sont des verbes hétéroclites au niveau exponence. Nous illustrons leurs propriétés en termes de schèmes flexionnels par les entrées de OPINOR « penser » et LOQUOR « parler » indiqués aux figures 7.11 et 7.12.

OPINOR					
I-PHON	opinor				
I-CAT	<i>verbe</i>				
MSF	{ <i>standard</i> }				
S-STEM	( <i>vide</i> )				
S-FORM	( <i>vide</i> )				
I-PAT		( $Z_{\text{REG-V-T}}^s, \text{id}$ ),	( $Z_a^{th}, \text{id}$ )	( $Z_{v-B}^{exp}, \text{id}$ )	( $Z_{adj-oa}^{exp}, \text{id}$ )
		( $Z_{\text{REG-V-T}}^s, \text{id}$ ),	( $Z_a^{th}, \text{id}$ )	( $z_{A,3}^{exp}, \text{id}$ )	( $Z_{adj-oa}^{exp}, \text{id}$ )
		( $Z_{\text{REG-V-T}}^s, \text{id}$ ),	( $Z_a^{th}, \text{id}$ )	( $z_{A,3}^{exp}, \text{id}$ )	( $Z_{adj-c}^{exp}, \text{id}$ )

FIGURE 7.11 : Analyse 3H : entrée lexicale du flexème OPINOR « penser »

12. La même chose vaut pour les passiva tantum.

LOQUOR					
I-PHON	loquor				
I-CAT	<i>verbe</i>				
MSF	{ <i>standard</i> }				
S-STEM	( <i>vide</i> )				
S-FORM	( <i>vide</i> )				
I-PAT		$(Z_{\text{REG-V-T}}^s, \text{id})$	$(Z_c^{th}, \text{id})$	$(Z_{v-D}^{exp}, \text{id})$	$(Z_{adj-oa}^{exp}, \text{id})$
		$(Z_{\text{REG-V-T}}^s, \text{id})$	$(Z_c^{th}, \text{id})$	$(z_{C,3}^{exp}, \text{id})$	$(Z_{adj-oa}^{exp}, \text{id})$
		$(Z_{\text{REG-V-T}}^s, \text{id})$	$(Z_c^{th}, \text{id})$	$(z_{C,3}^{exp}, \text{id})$	$(Z_{adj-c}^{exp}, \text{id})$

FIGURE 7.12 : Analyse 3H : entrée lexicale du flexème LOQUOR « parler »

### 7.1.6 Les schèmes flexionnels de l'analyse 3H

Compte tenu des combinaisons possibles entre les niveaux réalisationnels dans l'analyse 3H, notre description du système verbal latin repose sur l'existence de 87 schèmes flexionnels distincts. L'inventaire de ces schèmes avec le nombre de flexèmes associé à chacun d'entre eux est indiqué dans l'annexe B. L'implémentation en Alexina<sub>PTRESL</sub> emploie par ailleurs 36 variantes<sup>13</sup>. Le nombre de schèmes distincts variantes comprises est de 192.

## 7.2 Proposition d'une description à quatre niveaux de la flexion des verbes latins

Comme nous l'avons indiqué à la section 1.4.6 du chapitre 1, une critique possible de cette analyse<sup>14</sup> est qu'elle semble comporter une redondance non négligeable entre le niveau thématique et le niveau d'exponence. Elle nécessite notamment de stipuler quatre classes flexionnelles au niveau exposant. Or l'appartenance d'un flexème à l'une de ces classes est directement conditionnée par la nature de sa voyelle thématique : à savoir, les classes  $Z_{v-A}^{exp}$  et  $Z_{v-B}^{exp}$  pour les verbes appartenant aux thèmes  $Z_a^{th}$  et  $Z_e^{th}$  et les classes  $Z_{v-C}^{exp}$  et  $Z_{v-D}^{exp}$  pour les verbes appartenant aux thèmes  $Z_i^{th}$  et  $Z_c^{th}$  (y compris les

13. Le système des variantes en Alexina a été décrit à la section 5.2.1 page 285 du chapitre 5.

14. Nous remercions Grev Corbett et Matthew Baerman pour cette suggestion, lors d'une présentation au sein d'un atelier du SEMM.

flexèmes hétéroclites au niveau thématique). Ce fait se retrouve également dans les types d'hétéroclise observables pour les verbes déponents et les passiva tantum. On observera toujours des hétéroclites de type  $Z_{v-A}^{exp}$  et  $Z_{v-B}^{exp}$  ou  $Z_{v-C}^{exp}$  et  $Z_{v-D}^{exp}$ , mais jamais  $Z_{v-A}^{exp}$  et  $Z_{v-D}^{exp}$  ou  $Z_{v-C}^{exp}$  et  $Z_{v-B}^{exp}$ .

Dans la suite de ce chapitre nous proposons une analyse concurrente de celle du chapitre 1 et dont nous venons de présenter la formalisation, qui cherche à supprimer cette redondance. Cette nouvelle analyse, que nous appellerons désormais « analyse 4H », fait usage des quatre niveaux réalisationnels suivants <sup>15</sup>.

Niveau 1 niveau radical ;

Niveau 2 niveau thématique ;

Niveau 3 niveau exposant (exposants de temps, d'aspect et de mode) ;

Niveau 4 niveau exposant (marques de personne et de nombre).

À ces quatre niveaux s'ajoute, comme dans l'analyse 3H, un niveau pour la flexion casuelle des participes, empruntant les classes  $Z_{adj-c}^{exp}$  et  $Z_{adj-oa}^{exp}$  de la flexion adjectivale.

Une première différence entre les analyses 3H et 4H est que la répartition des exposants sur les deux niveaux (3 et 4) correspond à une différence entre les attributs réalisés à chacun d'entre eux : le temps, l'aspect et le mode (TAM) au niveau 3 et la personne, le nombre et au niveau 4.

### 7.2.1 Les classes réalisationnelles définies dans l'analyse 4H

L'analyse 4H ne pose à chaque niveau exposant que deux classes flexionnelles. Au niveau 3, la CLASSE I, notée  $Z_I^{exp}$ , correspond à l'inventaire des marques TAM comportant la marque *-bi-* pour le futur de l'indicatif, *-ba-* pour l'imparfait, *-nd* pour le gérondif et *-ns/-nd-* pour le participe présent. La classe CLASSE II, notée  $Z_{II}^{exp}$ , comporte l'inventaire des marques TAM comportant la marque *-e-* pour le futur de l'indicatif, *-eba-* pour l'imparfait et *-end* pour l'adjectif verbal.

Au niveau 4, la CLASSE A, notée  $Z_A^{exp}$ , est la classe des marques personnelles m-actives, la CLASSE B,  $Z_B^{exp}$ , la classe des marques m-passives.  $Z_A^{exp}$  ne comporte par ailleurs pas de règles réalisationnelles pour le participe passé et l'adjectif verbal, ainsi que pour l'infinitif futur qui se trouve cependant être réalisé par une périphrase. Il manque à  $Z_B^{exp}$  des exposants pour les participes futur et présent et le gérondif, ainsi que pour l'infinitif passé et pour toutes les formes perfectives finies qui sont par ailleurs périphrastiques.

15. L'implémentation complète de cette analyse en Alexina<sup>PARSL</sup> est disponible en ligne sur la page personnelle de l'auteur.



Nous résumons les quatre classes d'exponence avec leur nombre de flexèmes associés et leurs caractéristiques principales au tableau 7.9.

CLASSE	NOMBRE	SPÉCIFICITÉS	EXEMPLE	TRADUCTION
<b>NIVEAU 3</b>				
$Z_I^{exp}$	1 047	– PRS.IPFV.IND en <i>-ba-</i> , – FUT.IPFV en <i>-bi-</i> – GRV en <i>-nd</i> – PTCP.PRS en <i>-ns/-nd</i>	AMO AMOR	« <i>aimer</i> » « <i>être aimé</i> »
$Z_{II}^{exp}$	1 160	– PRS.IPFV.IND en <i>-eba-</i> , – FUT.IPFV en <i>-e- (-a-)</i> – GRV en <i>-end</i> – PTCP.PRS en <i>-ens/-end</i>	AUDIO AUDIOR	« <i>entendre</i> » « <i>être entendu</i> »
<b>NIVEAU 4</b>				
$Z_A^{exp}$	1 065	– désinences m-actives – pas de PTCP.PST, GRD – INF.FUT périphrastique	AMO AUDIO	« <i>aimer</i> » « <i>entendre</i> »
$Z_B^{exp}$	976	– désinences m-passives – pas de PTCP.PRS, GRV – INF.PST et PFV fini périphrastique	AMOR AUDIOR	« <i>être aimé</i> » « <i>être entendu</i> »
hétéroclites	166	$Z_A^{exp}$ et $Z_A^{exp}$	OPINOR LOQUOR	« <i>penser</i> » « <i>parler</i> »

TABLE 7.9 : Classes d'exponence définies au niveau 3 et 4 de l'analyse 4H

En ce qui concerne les niveaux thématique (2) et radical (1), ainsi que le niveau de la flexion casuelle participiale (5), l'analyse 4H pose exactement les mêmes classes réalisationnelles que l'analyse 3H. Dans la suite de ce chapitre, nous nous comparons les analyses 3H et 4H. Notons que ces deux analyses ne diffèrent que par leur grammaire flexionnelle. Le lexique flexionnel associé à chacune des grammaires est exactement le même.

Les figures 7.13 à 7.17 illustrent la représentation des entrées lexicales verbales latines dans notre nouvelle analyse 4H. Elles correspondent aux figures 7.2 à 7.5 représentant les entrées lexicales dans l'analyse concurrente 3H.

CONTINEO					
I-PHON	contineo				
I-CAT	<i>verbe</i>				
MSF	{ <i>standard</i> }				
S-STEM	( <i>vide</i> )				
S-FORM	( <i>vide</i> )				
I-PAT	$(Z_{\text{REG-U-T}}^s, \text{id}), (Z_e^{th}, \text{id}) (Z_I^{exp}, \text{id}) (Z_A^{exp}, \text{id}) (Z_{adj-oa}^{exp}, \text{id})$ $(Z_{\text{REG-U-T}}^s, \text{id}), (Z_e^{th}, \text{id}) (Z_I^{exp}, \text{id}) (Z_A^{exp}, \text{id}) (Z_{adj-c}^{exp}, \text{id})$				

FIGURE 7.13 : Analyse 4H : entrée lexicale du flexème CONTINEO « *contenir* »

CONTINEOR					
I-PHON	contineor				
I-CAT	<i>verbe</i>				
MSF	{ <i>standard</i> }				
S-STEM	( <i>vide</i> )				
S-FORM	( <i>vide</i> )				
I-PAT	$(Z_{\text{REG-U-T}}^s, \text{id}), (Z_e^{th}, \text{id}) (Z_I^{exp}, \text{id}) (Z_B^{exp}, \text{id}) (Z_{adj-oa}^{exp}, \text{id})$				

FIGURE 7.14 : Analyse 4H : entrée lexicale du flexème CONTINEOR « *être contenu* »

AUDIOR					
I-PHON	audior				
I-CAT	<i>verbe</i>				
MSF	{ <i>standard</i> }				
S-STEM	( <i>vide</i> )				
S-FORM	( <i>vide</i> )				
I-PAT	$(Z_{\text{REG-V-T}}^s, \text{id}), (Z_i^{th}, \text{id}) (Z_{II}^{exp}, \text{id}) (Z_B^{exp}, \text{id}) (Z_{adj-oa}^{exp}, \text{id})$				

FIGURE 7.15 : Analyse 4H : entrée lexicale du flexème AUDIOR « *être entendu* »

ACCIPIO					
I-PHON	accipio				
I-CAT	<i>verbe</i>				
MSF	{ <i>standard</i> }				
S-STEM	( <i>vide</i> )				
S-FORM	( <i>vide</i> )				
I-PAT	( $Z_{\text{REG-EQ-T}}^s$ , id),	( $z_{i,1}^{th}$ , id)	( $Z_{\text{II}}^{exp}$ , id)	( $Z_{\text{A}}^{exp}$ , id)	( $Z_{adj-oa}^{exp}$ , id)
	( $Z_{\text{REG-EQ-T}}^s$ , id),	( $z_{i,1}^{th}$ , id)	( $Z_{\text{II}}^{exp}$ , id)	( $Z_{\text{A}}^{exp}$ , id)	( $Z_{adj-c}^{exp}$ , id)
	( $Z_{\text{REG-EQ-T}}^s$ , id),	( $z_{c,2}^{th}$ , id)	( $Z_{\text{II}}^{exp}$ , id)	( $Z_{\text{A}}^{exp}$ , id)	( $Z_{adj-oa}^{exp}$ , id)

 FIGURE 7.16 : Analyse 4H : entrée lexicale du flexème ACCIPIO « *accepter* »

ACCIPIOR					
I-PHON	accipior				
I-CAT	<i>verbe</i>				
MSF	{ <i>standard</i> }				
S-STEM	( <i>vide</i> )				
S-FORM	( <i>vide</i> )				
I-PAT	( $Z_{\text{REG-EQ-T}}^s$ , id),	( $z_{i,1}^{th}$ , id)	( $Z_{\text{II}}^{exp}$ , id)	( $Z_{\text{B}}^{exp}$ , id)	( $Z_{adj-oa}^{exp}$ , id)
	( $Z_{\text{REG-EQ-T}}^s$ , id),	( $z_{c,2}^{th}$ , id)	( $Z_{\text{II}}^{exp}$ , id)	( $Z_{\text{B}}^{exp}$ , id)	( $Z_{adj-oa}^{exp}$ , id)

 FIGURE 7.17 : Analyse 4H : entrée lexicale du flexème ACCIPIOR « *être accepté* »

### 7.2.2 Interaction des classes réalisationnelles à travers les niveaux dans l'analyse 4H

Les classes réalisationnelles étant distribuée différemment dans 4H par rapport à 3H, elles s'interfaçent également de façon sensiblement différente dans les deux analyses. Dans 4H, les classes se combinent de façon croisée, comme indiqué à la figure 7.18.

On note que la CLASSE I ( $Z_{\text{I}}^{exp}$ ) est systématiquement associée aux classes thématiques en *a* ( $Z_{\text{a}}^{th}$ ) ou en *e* ( $Z_{\text{e}}^{th}$ ), tandis que la CLASSE II ( $Z_{\text{II}}^{exp}$ ) se combine à la classe thématique en *i*  $Z_{\text{i}}^{th}$  ou à la classe athématique  $Z_{\text{c}}^{th}$ . Au niveau 3, les classes d'exponence  $Z_{\text{I}}^{exp}$  et  $Z_{\text{II}}^{exp}$  factorisent ainsi cinq sous-classes de flexèmes classés selon leurs propriétés thématiques en

deux sous-classes seulement, en fonction de la réalisation de leurs propriétés de TAM. Les classes d'exponence du niveau 4 permettent de croiser les membres de ces deux sous-classes de niveau 3 en fonction de la réalisation m-active ou m-passive de leur propriétés de flexion en personne et nombre.

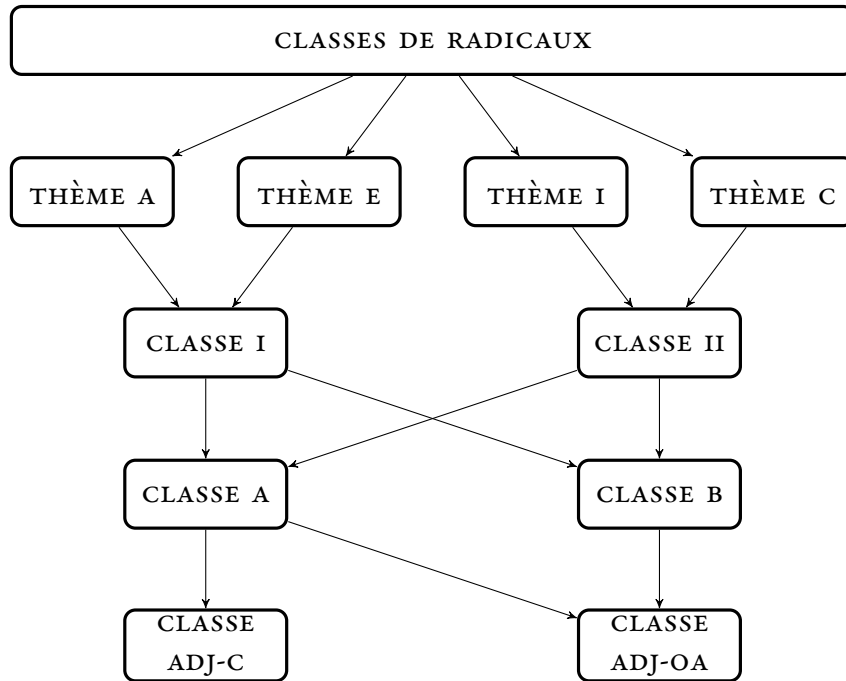


FIGURE 7.18 : Classes réalisationnelles dans l'analyse 4H

### 7.2.3 Représentation de la dérivation passive dans l'analyse 4H

Dans l'analyse 4H, les conséquences formelles de la dérivation passive sur la constitution des schèmes flexionnels se résument à l'association du flexème à la classe flexionnelle  $Z_A^{exp}$  (flexème de la base) ou à la classe flexionnelle  $Z_B^{exp}$  au niveau 4. Ce changement est illustré par les entrées des figures 7.13 et 7.14, ainsi que 7.16 et 7.17 : la classe  $Z_A^{exp}$ , classe du flexème de la base, y est indiquée en bleu, la classe  $Z_B^{exp}$  du dérivé correspondant en rouge.

L'avantage de cette analyse est qu'elle permet de rendre compte de la dérivation par un comportement systématique de changement de classe réalisationnelle  $Z_A^{exp}$  vers  $Z_B^{exp}$  au niveau 4 pour tous les flexèmes dérivables. En particulier, elle évite la mise en correspondance systématique de quatre classes flexionnelles deux à deux par la relation de dérivation, observée au niveau 3 de l'analyse 3H.

## 7.2.4 Modélisation des verbes déponents dans l'analyse 4H

OPINOR					
I-PHON	opinor				
I-CAT	verbe				
MSF	{ standard }				
S-STEM	(vide)				
S-FORM	(vide)				
I-PAT	( $Z_{\text{REG-V-T}}^s$ , id),	( $Z_a^h$ , id)	( $Z_I^{\text{exp}}$ , id)	( $Z_B^{\text{exp}}$ , id)	( $Z_{adj-oa}^{\text{exp}}$ , id)
	( $Z_{\text{REG-V-T}}^s$ , id),	( $Z_a^h$ , id)	( $Z_I^{\text{exp}}$ , id)	( $z_{A,3}^{\text{exp}}$ , id)	( $Z_{adj-oa}^{\text{exp}}$ , id)
	( $Z_{\text{REG-V-T}}^s$ , id),	( $Z_a^h$ , id)	( $Z_I^{\text{exp}}$ , id)	( $z_{A,3}^{\text{exp}}$ , id)	( $Z_{adj-c}^{\text{exp}}$ , id)

FIGURE 7.19 : Analyse 4H : entrée lexicale du flexème OPINOR « penser »

LOQUOR					
I-PHON	loquor				
I-CAT	verbe				
MSF	{ standard }				
S-STEM	(vide)				
S-FORM	(vide)				
I-PAT	( $Z_{\text{REG-V-T}}^s$ , id),	( $Z_a^h$ , id)	( $Z_{II}^{\text{exp}}$ , id)	( $Z_B^{\text{exp}}$ , id)	( $Z_{adj-oa}^{\text{exp}}$ , id)
	( $Z_{\text{REG-V-T}}^s$ , id),	( $Z_a^h$ , id)	( $Z_{II}^{\text{exp}}$ , id)	( $z_{A,3}^{\text{exp}}$ , id)	( $Z_{adj-oa}^{\text{exp}}$ , id)
	( $Z_{\text{REG-V-T}}^s$ , id),	( $Z_a^h$ , id)	( $Z_{II}^{\text{exp}}$ , id)	( $z_{A,3}^{\text{exp}}$ , id)	( $Z_{adj-c}^{\text{exp}}$ , id)

FIGURE 7.20 : Analyse 4H : entrée lexicale du flexème LOQUOR « parler »

Dans l'analyse 4H, les verbes déponents sont par conséquent uniquement hétéroclites au niveau 4<sup>16</sup>. Ils se fléchissent tous selon des zones des classes  $Z_A^{\text{exp}}$  et  $Z_B^{\text{exp}}$ . En particulier, il n'y a de ce point de vue plus de différence entre les verbes appartenant aux classes thématiques  $Z_a^h$  et  $Z_e^h$  comme OPINOR « penser » d'une part et  $Z_c^h$  et  $Z_i^h$  comme LOQUOR « parler »

16. Sauf cas d'hétéroclise thématique pour les verbes comme ACCIPIO « accepter ».

de l'autre. Nous indiquons des exemples de ces deux types d'entrées déponentes pour les verbes *OPINOR* et *LOQUOR* aux figures 7.19 et 7.20 respectivement.

## 7.3 Mesure de la longueur de description des analyses concurrentes

### 7.3.1 Comparaison de la longueur de description des analyses 3H et 4H

En suivant les formalisations décrites ci-dessus, nous avons implémenté les deux descriptions concurrentes 3H et 4H de la flexion des verbes latins en *Alexina*<sub>PARSLI</sub>. Les deux descriptions sont associées au même ensemble de 2 271 verbes et comportent le même nombre de schèmes flexionnels, de classes de thèmes et de radicaux, d'opérations morphologiques définies, de règles morphologiques et le même inventaire de traits morphosyntaxiques. La seule différence entre les deux analyses relève de la structuration des niveaux d'exponence.

Nous avons mesuré la longueur de description de chacune des analyses à l'aide de la mesure de compacité dédiée présentée au chapitre 6. Les résultats des deux mesures sont indiqués au tableau 7.10 et illustrés par le diagramme de Kiviat de la figure 7.21.

Les résultats montrent que les deux descriptions ont une différence de compacité globale très faible. Quand bien même la description à quatre niveaux réalisationnels (en bleu) obtient pour sa LDM globale à environ 65,6 kilobits un score légèrement meilleur que celle de la description à trois niveaux (en rouge) à environ 65,9 kilobits, la différence de 0,3 kilobits, soit environ 0,5% des descriptions respectives, n'est pas suffisante pour pouvoir classer en termes de compacité les deux descriptions. En termes de compacité globale, les deux descriptions sont ainsi sensiblement équivalentes.

Ce résultat n'est cependant pas surprenant : les deux analyses sont conçues pour être identiques en tout point de vue excepté le traitement de la structure du niveau d'exponence. Or, comme le montre très clairement le diagramme de Kiviat de la figure 7.21, la plus grande partie de l'information des deux description est comprise dans le lexique, et notamment dans la part de formes et radicaux supplétifs qu'il comprend. Dans la mesure où la différence entre les deux descriptions se cantonne à la structuration de la grammaire flexionnelle, comparer les deux descriptions globalement a pour conséquence de noyer la différence de taille de la grammaire dans la quantité d'information contenue dans le lexique. En ne considérant que la seule grammaire, la différence entre les deux descriptions

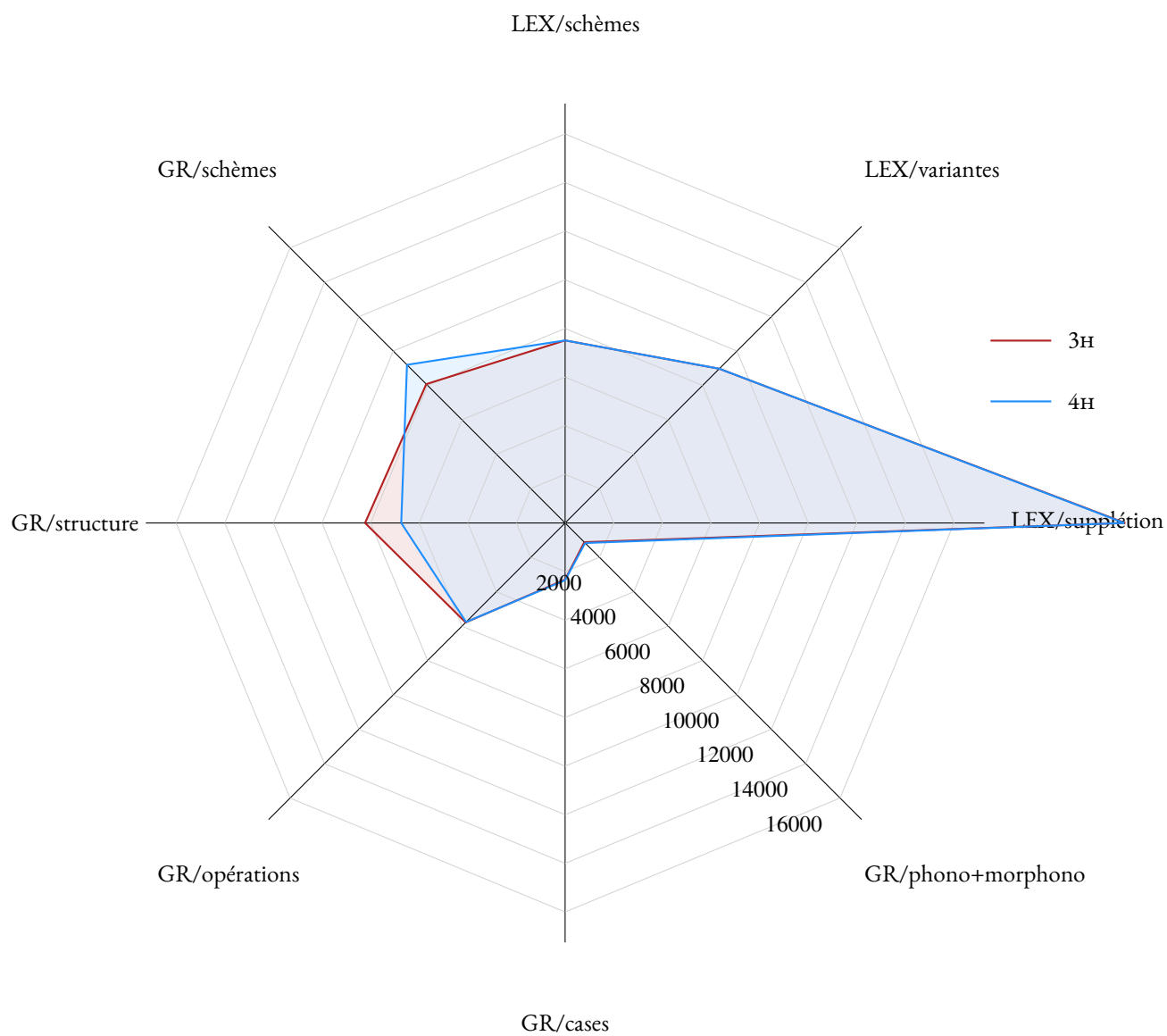


FIGURE 7.21 : Mesures de compacité globale (lexique + grammaire)s des deux descriptions de la flexion verbale latine

	DESCR. À 3 NIVEAUX (3H)	DESCR. À 4 NIVEAUX (4H)
<b>LEXIQUE</b>	40 369	40 333
schèmes	7 509	7 509
variantes	8 970	8 973
supplétion	23 044	23 005
<b>GRAMMAIRE</b>	25 538	25 225
définition des cases	2 327	2 354
schèmes	8 077	9 200
structure	8 229	6 742
opérations	5 786	5 762
(mor-)phonologie	1 118	1 058
<b>TOTAL</b>	65 907	65 558

TABLE 7.10 : Résultats des mesures de compacité sur les deux descriptions de la flexion verbale latine (en bits)

correspond ainsi à une différence d'environ 1,2% de la grammaire 3H. À la figure 7.22, nous illustrons ainsi la différence pour les mesures sur la seule grammaire dans un second diagramme de Kiviat. La différence reste néanmoins faible.

L'avantage de la mesure de compacité que nous avons présentée au chapitre 6 est qu'elle a été conçue pour pouvoir identifier, au sein d'une description, la part d'information contenue dans chacune de ses composantes principales. Nous pouvons ainsi mesurer spécifiquement la différence de compacité dans les composantes de la grammaire, et au sein de la grammaire dans la partie de la description qui diffère entre 3H et 4H et évaluer la différence en termes de compacité qu'introduisent les variations entre la description 3H et la description 4H.

Le second diagramme de Kiviat montre très clairement que la partie différente entre les deux analyses concerne la STRUCTURE générale de la grammaire (en termes de nombre de niveaux et de des classes réalisationnelles au niveau exponence) ainsi que la définition des SCHÈMES flexionnels.

La partie STRUCTURE comporte notamment la définition des classes flexionnelles. Il



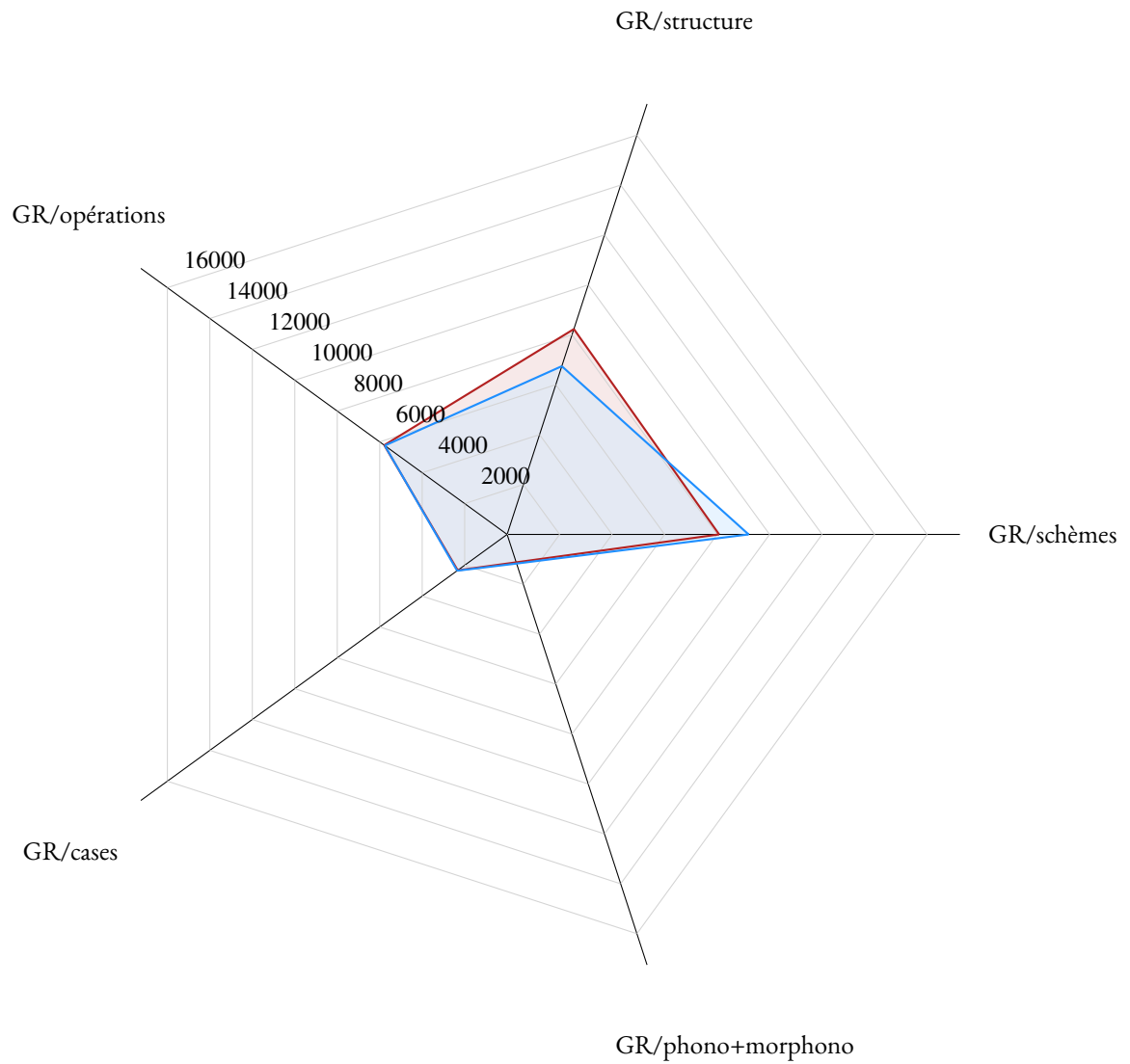


FIGURE 7.22 : Mesures de compacité pour la seule grammaire flexionnelle des deux descriptions de la flexion verbale latine

s'agit donc exactement de la composante que nous avons fait varier afin de répondre à la critique de redondance en termes de classes flexionnelles faites à la description 3H. Pour la composante STRUCTURE nos deux descriptions donnent ainsi les mesures suivantes :

**Structure** Environ 8,2 kilobits pour 3H et 6,7 kilobits pour 4H, soit une différence d'environ 1,5 kilobitsbit correspondant à 18,1% de la part d'information comprise dans la structure de 3H en faveur de 4H.

Grâce à notre mesure plus précise, nous pouvons ainsi constater que supprimer dans la composante STRUCTURE la redondance de l'analyse 3H critiquée pour des raisons théoriques mène effectivement à une amélioration notable de la compacité dans la partie de la description concernée par cette critique.

En revanche, nous observons dans le diagramme de la figure 7.22 que la composante STRUCTURE n'est pas la seule à varier entre les deux descriptions. En effet, nous observons également une différence notable de longueur de description dans la composante SCHÈMES. Cette différence est cette fois en faveur de la description 3H. Nous donnons ci-dessous les valeurs de compacité de la composante SCHÈMES.

**schèmes** 8,1 kilobits pour 3H et 9,2 kilobits pour 4H, soit une différence d'environ 1,5 kilobit correspondant à 13,8% de la part d'information comprise dans les schèmes de 3H en faveur de 3H.

Cette différence est également assez importante. La complexité comprise dans la définition des classes flexionnelles a ainsi été pour une grande partie déplacée vers la définition des schèmes flexionnels dans la grammaire.

Nous constatons donc pour les deux composantes intervenant dans la différence entre les deux analyse

3H 8,1 kilobits pour les schèmes et 8,2 kilobits pour la structure, soit un total de 16,3 bits.

4H 9,2 kilobits pour les schèmes et 6,7 kilobits pour la structure, soit un total de 15,9 kilobits.

La différence de longueur de description sur la partie différente des deux descriptions revient ainsi à 2,23% de la longueur de description de ces deux composantes dans la description 3H. Cette différence est plus importante que la différence globale entre les deux descriptions, mais elle reste néanmoins faible.

Nous constatons ainsi à l'aide de notre mesure de compacité que même avec des analyses conceptuellement différentes, la différence en termes de longueur de description peut être faible. Compte tenu de la redistribution nécessaire au moment de la conception de la

nouvelle analyse, l'analyse 3H, critiquée pour sa redondance dans la définition des classes flexionnelles  $Z_{v-A}^{exp}$ ,  $Z_{v-B}^{exp}$ ,  $Z_{v-C}^{exp}$  et  $Z_{v-D}^{exp}$  est en fin de compte à peine moins compacte que l'analyse 4H, et ce quand bien même on ne mesure que les parties entre les descriptions qui sont réellement différentes.

En revanche, nous avons pu observer des différences locales entre les composantes des deux descriptions. Dans la mesure où la différence rapportée à l'intégralité de la description est négligeable, une simple compression des deux descriptions n'aurait pas été capable de montrer la différence réelle observable dans les sous-parties modifiées des deux descriptions. En particulier, l'amélioration qu'apporte 4H en termes de compacité dans la partie structure serait passée complètement inaperçue compte tenu du fait que la dégradation de la compacité dans la composante des SCHÈMES contrebalance presque entièrement cette amélioration. À défaut de pouvoir nous aider à trancher clairement en faveur de l'une ou de l'autre analyse, notre mesure nous aura, dans ce cas, permis de saisir quantitativement pourquoi la seconde analyse ne diffère en fin de compte pas grandement de la première.

**Analyses concurrentes avec hétéroclise au niveau radical** Compte tenu du grand nombre de zones de radicaux partagées entre les différentes classes de radicaux, nous avons également implémenté deux analyses concurrentes pour 3H et 4H dans lesquelles toutes les classes de radicaux du tableau 7.8 commençant par H- ont été supprimées. Les flexèmes leur étant associés ont été traités comme des flexèmes hétéroclites au niveau radical. Nous illustrons ce cas sur l'entrée de GERO « tourner » aux figures 7.23 et 7.24<sup>17</sup>. GERO est indiqué comme ayant un schème flexionnel de type  $\mathbb{P}_{v-cC-H-SS-T}$  : il emploie donc la classe thématique  $Z_c^{th}$ , la classe d'exponence  $Z_{v-C}^{exp}$  en 3H<sup>18</sup> et, dans les analyses de base, la classe de radicaux  $Z_{H-SS-T}^s$ . L'entrée 7.23 correspond à l'entrée sans hétéroclise radicale de nos analyses. L'entrée 7.24 correspond à la version concurrente où GERO est hétéroclite au niveau radical et emploie la zone  $z_{BASE,1}^s$  pour construire son RADICAL 1, la  $z_{SS,2}^s$  pour construire son RADICAL 2 et la zone  $z_{T,3}^s$  pour construire son RADICAL 3.

Pour 3H comme 4H, l'analyse concurrente avec hétéroclise diminue le nombre de classes de radicaux : elles ne sont plus qu'au nombre de huit au lieu de 17, mais elle rend considérablement plus complexes les schèmes flexionnels des entrées hétéroclites. Le gain en termes de classes de radicaux moins nombreuses n'est en fin de compte pas suffisant pour contrebalancer la perte de compacité due aux schèmes plus complexes. Les valeurs associées

17. Les exemples sont indiqués avec les entrées correspondant à l'analyse 3H.

18.  $Z_{II}^{exp} + Z_A^{exp}$  en 4H.

GERO					
I-PHON	gero				
I-CAT	<i>verbe</i>				
MSF	{ <i>standard</i> }				
S-STEM	( <i>vide</i> )				
S-FORM	( <i>vide</i> )				
I-PAT		$(Z_{H-SS-T}^s, id)$ ,	$(Z_c^{th}, id)$	$(Z_{v-C}^{exp}, id)$	$(Z_{adj-oa}^{exp}, id)$
		$(Z_{H-SS-T}^s, id)$ ,	$(Z_c^{th}, id)$	$(Z_{v-C}^{exp}, id)$	$(Z_{adj-c}^{exp}, id)$

FIGURE 7.23 : Analyse 3H : Entrée lexicale du flexème de la base GERO « tourner » sans hétéroclise au niveau radical

GERO					
I-PHON	gero				
I-CAT	<i>verbe</i>				
MSF	{ <i>standard</i> }				
S-STEM	( <i>vide</i> )				
S-FORM	( <i>vide</i> )				
I-PAT		$(z_{BASE,1}^s, id)$ ,	$(Z_c^{th}, id)$	$(Z_{v-C}^{exp}, id)$	$(Z_{adj-c}^{exp}, id)$
		$(z_{BASE,1}^s, id)$ ,	$(Z_c^{th}, id)$	$(Z_{v-C}^{exp}, id)$	$(Z_{adj-oa}^{exp}, id)$
		$(z_{SS,2}^s, id)$ ,	$(Z_c^{th}, id)$	$(Z_{v-C}^{exp}, id)$	
		$(Z_{T,3}^s, id)$ ,	$(Z_c^{th}, id)$	$(Z_{v-C}^{exp}, id)$	$(Z_{adj-oa}^{exp}, id)$

FIGURE 7.24 : Analyse 3H : Entrée lexicale du flexème de la base GERO « tourner » avec hétéroclise au niveau radical

aux analyses avec et sans hétéroclise radicale pour 3H sont ainsi :

- 65,9 kilobits pour l’analyse sans hétéroclise radicale dans 3H, dont 25,5 kilobits pour la seule grammaire flexionnelle ;
- 66,2 kilobits pour l’analyse avec hétéroclise radicale dans 3H, dont 25,8 pour la seule grammaire flexionnelle.

Sans surprise, les différences entre 4H avec hétéroclise et 4H sans hétéroclise correspondent au rapport observé entre 3H avec hétéroclise et 3H sans hétéroclise.

Des deux types de descriptions, l’analyse de départ (sans hétéroclise radicale) apparaît comme *très* légèrement plus compacte : on mesure environ 1% de longueur de description en plus pour les analyses avec hétéroclise radicale. Bien qu’elles ne présentent qu’un avantage négligeable en termes de compacité, c’est avec les analyses sans hétéroclise radicale que nous comparerons une troisième analyse concurrente.

### 7.3.2 Comparaison de la longueur de description de nos analyses avec l’analyse traditionnelle

Au chapitre 1, nous avons opposé notre analyse de la flexion verbale latine à l’analyse traditionnelle dans laquelle la voix est considérée comme un attribut de la flexion. Nous avons donné des arguments théoriques en faveur d’une analyse dérivationnelle.

Nous n’avons cependant pas encore donné d’évaluation quantitative en termes de compacité de la différence entre l’analyse traditionnelle et la nôtre. Afin de pouvoir évaluer la différence de longueur de description entre notre analyse et celle d’une description comparable de l’analyse traditionnelle, nous avons également, à partir des données de *Leffa*, reconstitué une description de la flexion verbale latine dans laquelle le changement de voix relèverait de la flexion<sup>19</sup>. Cette analyse est associée à son propre lexique flexionnel *Leffa*<sub>TRAD</sub>. Nous esquissons ici rapidement les caractéristiques formelles de la modélisation en  $\mathbb{P}\mathbb{R}\mathbb{S}\mathbb{L}$  de cette description et en mesurons la compacité pour comparaison avec notre propre analyse.

Pour développer la description correspondant à l’analyse traditionnelle, que nous appellerons désormais 3F<sup>20</sup>, nous sommes partis de la description 3H. Les points communs entre les analyses 3F et 3H sont les suivants :

---

19. Ici, les tableaux 7.1 à 7.4 correspondent ainsi chacun au paradigme type d’un verbe unique.

20. En raison du nombre de niveaux réalisational (3) et du fait que le passif y est traité comme de la flexion.

- Les deux descriptions sont structurées en trois niveaux réalisationnels proprement verbaux et un niveau 4 pour la flexion casuelles des seuls participes, gérondif et supin.
- Les niveaux radical et thématique sont identiques entre les deux descriptions<sup>21</sup>.
- Les règles morphologiques sont identiques.

La nouvelle analyse repose cependant sur la définition d'une catégorie flexionnelle verbale différente de celle définie pour 3H. Dans la mesure où, dans cette analyse, les verbes intransitifs ont un paradigme deux fois plus petit que les verbes transitifs, nous aurions pu poser des catégories flexionnelles différentes pour ces deux sous-ensembles de verbes. Cette question n'est cependant généralement pas explicitement posée dans la littérature. Afin de ne pas introduire plus de différences que nécessaire entre les analyses, nous avons pris le parti de ne poser qu'une seule catégorie flexionnelle en 3F comme en 3H. Cette catégorie flexionnelle diffère néanmoins dans sa définition de celle définie en 3H. Chaque flexème verbal y est défini comme réalisant en tout 263 formes<sup>22</sup>, dont 60 formes finies actives et autant de passives, 11 formes impératives, trois infinitifs, 3×30 formes participiales, 30 formes pour l'adjectif verbal, cinq formes pour le gérondif et deux pour le supin. Nous indiquons les propriétés flexionnelles de la catégorie flexionnelle définie en 3F à la figure 7.25. Les différences majeures par rapport à la catégorie flexionnelle de 3H sont indiquées en rouge.

Dans cette analyse, tout verbe intransitif non déponent a ainsi un paradigme déficient pour les formes passives, le participe passé, l'adjectif verbal et le supin. Les verbes déponents et les activa tantum sont également déficients, mais moins que les verbes transitifs normaux dans la mesure où il ne leur manque que les formes actives finies et l'infinitif passif. Les verbes déponents présentent en revanche de surcroît des cas de décalage morphosyntaxique dont la fonction de transfert (utilisation des formes passives pour exprimer l'actif) doit être définie dans la grammaire.

En ce qui concerne les classes flexionnelles, l'analyse traditionnelle est plus simple que l'analyse 3H dans la mesure où elle n'a besoin de spécifier que deux classes, puisque les formes actives et passives appartiennent canoniquement à un même paradigme. Les deux classes flexionnelles définies dans 3F sont ainsi la classe flexionnelle avec l'indicatif du futur en *-bi* et de l'imparfait en *-ba*<sup>23</sup> que nous noterons  $Z_{v-1}^{exp}$  et la classe réalisant du futur en

21. Nous aurions pu définir un thème spécifique aux verbes thématiquement hétéroclites comme ACCIPIO « *accepter* », mais compte tenu du fait que de nombreuses analyses traditionnelles considèrent ce thème comme un thème mixte, donc implicitement hétéroclite, nous avons pris le parti de ne maintenir que quatre classes thématiques et de continuer à traiter les verbes comme ACCIPIO comme des cas d'hétéroclise.

22. Au lieu de 196 pour les analyses 3H et 4H.

23. Le participe présent en *-ns/-nd* et le gérondif et l'adjectif verbal en *-nd*.

VERBE	
I-CAT	<i>verbe</i>
MSF	$\left\{ \begin{array}{l} (1\vee2\vee3).(SG\vee PL).(PRS\vee PST).(IPFV\vee PFV).(IND\vee SBJV).(ACT\vee PASS), \\ (1\vee2\vee3).(SG\vee PL).FUT.(IPFV\vee PFV).IND.(ACT\vee PASS), \\ 2.(SG\vee PL).PRS.IMP.(ACT\vee PASS), (2\vee3).SG.FUT.IMP.(ACT\vee PASS), \\ (2\vee3).SG.FUT.IMP.ACT, 3.FUT.PL.PASS, (PRS\vee PST).INF.(ACT\vee PASS), \\ INF.FUT.ACT, GRV.(NOM\vee ACC\vee GEN\vee DAT\vee ABL), \\ (PRS\vee PST\vee FUT).PTCP.(NOM\vee ACC\vee GEN\vee DAT\vee ABL).(SG\vee PL).(M\vee F\vee N), \\ GRD.(NOM\vee ACC\vee GEN\vee DAT\vee ABL).(SG\vee PL).(M\vee F\vee N), SUP.(NOM\vee ACC).SG \end{array} \right\}$
I-PAT	$\begin{array}{l} (Z^s, id), (Z^{th}, id), (Z^{exp}, id), (Z_{adj-oa}^{exp}, id) \\ (Z^s, id), (Z^{th}, id), (Z^{exp}, id), (Z_{adj-c}^{exp}, id) \end{array}$

FIGURE 7.25 : 3F : catégorie flexionnelle verbale en latin

$-e^{24}$  et de l'imparfait en  $-eba^{25}$  que nous noterons  $Z_{v-2}^{exp}$ .

Nous indiquons aux figures 7.26 à 7.28 trois entrées types de cette analyse : l'entrée d'AMO « *aimer* » pour illustrer un verbe transitif régulier, l'entrée de LUGEO « *briller* » pour un verbe intransitif et LOQUOR « *parler* » pour illustrer une entrée déponente.

L'entrée de AMO est parfaitement régulière.

AMO	
I-PHON	<i>amo</i>
I-CAT	<i>verbe</i>
MSF	{ <i>standard</i> }
S-STEM	( <i>vide</i> )
S-FORM	( <i>vide</i> )
I-PAT	$\begin{array}{l} (Z_{REG-V-T}^s, id), (Z_a^{th}, id), (Z_{v-1}^{exp}, id), (Z_{adj-oa}^{exp}, id) \\ (Z_{REG-V-T}^s, id), (Z_a^{th}, id), (Z_{v-1}^{exp}, id), (Z_{adj-c}^{exp}, id) \end{array}$

 FIGURE 7.26 : Analyse 3F : entrée lexicale du flexème AMO « *aimer* »

L'entrée de LUGEO est déficiente : il lui manque les formes finies passives, dont le PTCP.PST et le GRD.

24. En  $-a$  pour le 1SG.FUT.IPFV.IND.

25. Le participe présent en  $-ens/-end$  et le gérondif et l'adjectif verbal en  $-end$ .

LUGEO	
I-PHON	lugeo
I-CAT	verbe
MSF	$\left\{ \begin{array}{l} (1\vee2\vee3).(SG\vee PL).(PRS\vee PST).(IPFV\vee PFV).(IND\vee SBJV).ACT, \\ (1\vee2\vee3).(SG\vee PL).FUT.(IPFV\vee PFV).IND.ACT, \\ 2.(SG\vee PL).PRS.IMP.ACT, (2\vee3).SG.FUT.IMP.ACT, \\ (2\vee3).SG.FUT.IMP.ACT, (PRS\vee PST\vee FUT).INF.ACT, \\ GRV.(NOM\vee ACC\vee GEN\vee DAT\vee ABL), \\ (PRS\vee FUT).PTCP.(NOM\vee ACC\vee GEN\vee DAT\vee ABL).(SG\vee PL).(M\vee F\vee N), \\ SUP.(NOM\vee ACC).SG \end{array} \right\}$
S-STEM	(vide)
S-FORM	(vide)
I-PAT	$\begin{array}{cccc} (Z_{REG-X-T}^s, id), & (Z_a^{th}, id) & (Z_{v-1}^{exp}, id) & (Z_{adj-oa}^{exp}, id) \\ (Z_{REG-X-T}^s, id), & (Z_a^{th}, id) & (Z_{v-1}^{exp}, id) & (Z_{adj-c}^{exp}, id) \end{array}$

FIGURE 7.27 : Analyse 3F : entrée lexicale du flexème LUGEO « briller »

L'entrée de LOQUOR est également déficiente : il lui manque les formes finies passives, mais pas le PTCP.PST et le GRD. Les cases du paradigme de LOQUOR sont réalisées par l'emploi des seules zones flexionnelles  $z_{v-1,pass}^{exp}$  des formes passives (dont PTCP.PST et GRD) et  $z_{v-1,NFa}^{exp}$  des formes non finies actives PTCP.PRS, PTCP.FUT, GRV et SUP. En plus d'être déficient, LOQUOR est ainsi également défectif. Pour la réalisation de ses formes actives finies, elle possède par ailleurs une fonction de transfert  $t_{depo}$  au niveau 3 qui indique que ces formes actives sont réalisées par la zone de flexion  $z_{v-2,pass}^{exp}$  pour le flexème LOQUOR, mais que la forme produite porte néanmoins bien la valeur *active*. En d'autres termes, un verbe déponent présente un cas de décalage morphosyntaxique au niveau 3.

Enfin, une différence importante entre les deux analyses concerne le lexique flexionnel. Dans le lexique flexionnel de 3F, les entrées passives dérivées d'entrées actives de *Leffla* ont été fusionnées avec les entrées actives. Le lexique flexionnel *Leffla*<sub>TRAD</sub> ne comporte ainsi plus que 1 276 entrées, mais il produit les mêmes 263 677 formes fléchies.

Les mesures de compacité appliquées à l'analyse 3F sont résumées dans la troisième colonne de chiffres du tableau 7.11, à côté des mesures des deux autres descriptions. Nous illustrons également les trois ensembles de mesures dans le diagramme de Kiviat de la figure 7.29. L'analyse 3F y est indiquée en vert, 3H en rouge et 4H en bleu.

Le diagramme de Kiviat montre très clairement que la description 3F inspirée des



LOQUOR	
I-PHON	loquor
I-CAT	verbe
MSF	$\left\{ \begin{array}{l} (1\vee2\vee3).(SG\vee PL).(PRS\vee PST).(IPFV\vee PFV).(IND\vee SBJV).ACT, \\ (1\vee2\vee3).(SG\vee PL).FUT.(IPFV\vee PFV).IND.ACT, \\ 2.(SG\vee PL).PRS.IMP.(ACT), (2\vee3).SG.FUT.IMP.ACT, \\ , (PRS\vee PST).INF.ACT \\ GRV.(NOM\vee ACC\vee GEN\vee DAT\vee ABL) \\ (PRS\vee PST\vee FUT).PTCP.(NOM\vee ACC\vee GEN\vee DAT\vee ABL).(SG\vee PL).(M\vee F\vee N) \\ GRD.(NOM\vee ACC\vee GEN\vee DAT\vee ABL).(SG\vee PL).(M\vee F\vee N), SUP.(NOM\vee ACC).SG \end{array} \right\}$
S-STEM	(vide)
S-FORM	(vide)
I-PAT	$\begin{array}{l} (Z_{REG-V-T}^S, id), (Z_a^{th}, id) (z_{v-1,pass}^{exp}, t_{depo}) (Z_{adj-oa}^{exp}, id) \\ (Z_{REG-V-T}^S, id), (Z_a^{th}, id) (z_{v-1,NFa}^{exp}, id) (Z_{adj-c}^{exp}, id) \end{array}$

FIGURE 7.28 : Analyse 3F : entrée lexicale du flexème LOQUOR « briller »

descriptions traditionnelles est bien plus compacte que les nôtres. En particulier la taille du lexique et des schèmes flexionnels est nettement réduite par rapport aux deux autres analyses. La différence entre 3H et 3F vaut ainsi environ 7,3 kilobits, soit 11,5% de la longueur de la description 3H. Dans la mesure où notre mesure de compacité fait abstraction de l'identité des entrées lexicales, elle ne prend ici pas en compte la longueur de description des formes de citation en tant que telles. Cette différence était nulle dans la comparaison entre 3H et 4H, le lexique étant le même pour les deux. En plus de comporter moins de schèmes flexionnels, 3F comporte en revanche aussi moins d'entrées lexicales que 3H/4H. Si la longueur de description des formes de citation était prise en compte, la différence de compacité entre 3F d'une part et 3H/4H de l'autre n'en serait que plus marquée encore.

Des considérations théoriques extérieures à la flexion nous avaient poussée à proposer l'analyse dérivationnelle du passif en latin et de concevoir en conséquence l'analyse 3H. Les mesures de compacité montrent que ces choix théoriques argumentés ne mènent pas forcément à la description la plus compacte. Toutefois, nous n'avons dans les trois analyses formalisé, implémenté et mesuré que la composante flexionnelle des trois analyses. Or les considérations qui nous avaient amenée à proposer 3H ne relevaient pas de propriétés internes au système flexionnel. Choisir ici la description la plus compacte sur de purs

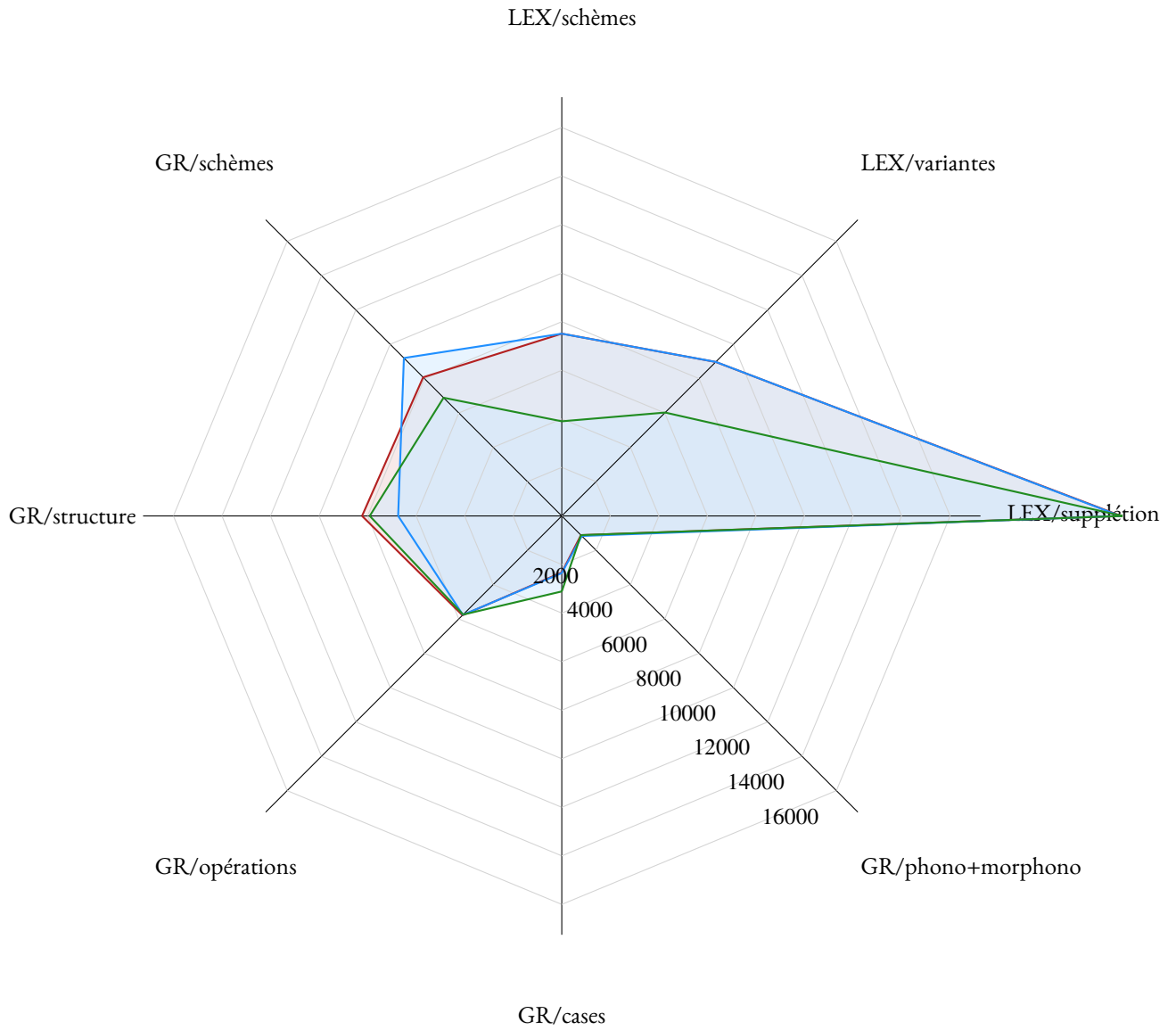


FIGURE 7.29 : Mesures de compacité globale (lexique + grammaire)s des trois descriptions de la flexion verbale latine

	3H	4H	3F
<b>LEXIQUE</b>	40 369	40 333	33 848
schèmes	7 509	7 509	3 899
variantes	8 970	8 973	6 024
supplétion	23 044	23 005	23 103
<b>GRAMMAIRE</b>	25 538	25 225	24 768
définition des cases	2 327	2 354	3 109
schèmes	8 077	9 200	6 887
structure	8 229	6 742	7 912
opérations	5 786	5 762	5 745
(mor-)phonologie	1 118	1 058	1 116
<b>TOTAL</b>	65 907	65 558	58 616

TABLE 7.11 : Résultats des mesures de compacité sur les trois descriptions de la flexion verbale latine (en bits)

critères flexionnels n'aurait par conséquent que peu de sens. Afin de véritablement pouvoir mesurer la différence entre ces deux descriptions, il faudrait également pouvoir rendre compte dans la modélisation, l'implémentation et les mesures des propriétés dérivationnelles, et notamment sémantiques, des trois descriptions. Les mesures auxquelles nous sommes confrontées ici montrent ainsi la limite de notre métrique : il s'agit d'une métrique de la flexion en tant que composante autonome. Elle ne saurait être utilisée pour comparer deux descriptions qui ne sont pas comparables de par leurs propriétés externes trop différentes. En particulier, le résultat plus compact de l'analyse 3F ne constitue en rien un critère suffisant pour la préférer aux deux autres.

## 7.4 Mesures de canonicité

À la section précédente, nous avons constaté que nos deux analyses supposant la passivation comme un cas de dérivation se valent en termes de compacité. Nous n'avons donc pas de raison forte pour préférer, de ce point de vue, l'une des deux analyses plutôt que l'autre.

Nous avons également constaté que l'analyse 3F inspirée des analyses traditionnelles et plus compacte que les deux autres peut difficilement être préférée à ces dernières sur le seul critère de sa compacité plus grande, les présupposés théoriques extérieurs à la flexion étant trop différents entre les deux types de représentation.

Une autre façon dont nous pouvons cependant comparer nos descriptions consiste à mesurer le degré de canonicité qu'elles comportent pour chacun des phénomènes non canoniques que nous pourrions identifier. Notons qu'en termes de canonicité, nos deux analyses présentent exactement les mêmes phénomènes. Les mesures que nous allons appliquer dans cette section seront donc identiques pour l'analyse 3H et pour l'analyse 4H. Nous comparerons donc directement les indices de canonicités communs à ces deux descriptions aux indices de canonicité de l'analyse 3F plus traditionnelle.

### 7.4.1 Identification des phénomènes non canoniques présents

Les phénomènes non canoniques observés sont les mêmes dans les descriptions 3H et 4H, mais différent pour l'analyse 3F.

**Hétéroclise** Les deux descriptions font état de cas d'hétéroclise à divers niveaux.

**3H : exposants** La description 3H présente de l'hétéroclise au niveau exposant (niveau 3) entre les classes flexionnelles  $Z_A^{exp}$  et  $Z_B^{exp}$  d'une part et les classes flexionnelles  $Z_C^{exp}$  et  $Z_D^{exp}$  de l'autre ; ces cas d'hétéroclise concernent autant les verbes déponents que les passiva tantum. Ce type d'hétéroclise concerne en tout 166 flexèmes.

**4H : exposants** Pour ces mêmes 166 verbes, la description 4H présente des cas d'hétéroclise au second niveau exposant (niveau 4) entre les classes flexionnelles  $Z_A^{exp}$  et  $Z_B^{exp}$ .

**3F : exposants** Dans la mesure où les classes flexionnelles y couvrent les formes m-actives et m-passives, la description 3F ne comporte pas de cas d'hétéroclise au niveau exponence.

**Thèmes** Les trois descriptions présentent des cas d'hétéroclise thématique au niveau 2 pour les verbes à thème mixte. Ces cas d'hétéroclise concernent ainsi la classes thématique  $Z_i^{th}$  et la classe athématique  $Z_C^{th}$ . Les verbes présentant de l'hétéroclise thématique sont au nombre de 126 dans 3H et 4H. Dans 3F, on en compte 66.

**Déficiência** On observe des instances de déficiência dans les trois descriptions.

**3H/4H** 3H et 4H présentent treize cas de déficience régulière. Ces cas sont définis dans la grammaire et associés à 64 entrées lexicales.

**3F** Dans 3F, les verbes intransitifs, les verbes déponents et les activa tantum sont systématiquement déficients. 3F comporte en tout 21 types de déficience. Elle concerne 154 entrées, dont 95 verbes intransitifs réguliers.

**Supplétion** On trouve des cas de supplétion de radicaux autant que de supplétion de formes.

**Radicaux 3H/4H** 184 entrées présentent de la supplétion de radicaux pour 377 radicaux supplétifs indiqués dans le lexique flexionnel *Leffa*.

**3F** Dans 3F, 109 flexèmes ont des radicaux supplétifs. En tout *Leffa*<sub>TRAD</sub> en comporte 229.

**Forme 3H/4H** La supplétion de formes concerne 22 entrées lexicales. Le lexique flexionnel *Leffa* comporte en tout 410 formes supplétives.

**3F** Dans 3F, la supplétion de formes concerne également 22 entrées lexicales et le nombre de formes supplétives dans *Leffa*<sub>TRAD</sub> est de 410, c'est-à-dire égal au nombre dans *Leffa*.

**Allomorphie radicale** L'allomorphie radicale est un phénomène généralisé touchant toutes les entrées. Chacune<sup>26</sup> possède soit deux soit trois radicaux différents formés régulièrement à partir de classes de radicaux définies dans la grammaire.

**Défectivité** On observe des cas de défectivité au sens strict, mais également des cas de réalisation périphrastique. Ces deux phénomènes sont, à ce stade, représentés de la même manière en  $\mathbb{P}\mathbb{R}\mathbb{S}\mathbb{L}$ .

**Cas systématiques** Les cas de véritable défectivité :

**3H/4H** Les verbes présentent selon leur classe d'exponence, un certain nombre de cas de défectivité systématique. Il n'existe notamment pour les verbes à désinence m-active pas de participes passé ni d'adjectif verbal ; inversement les verbes à désinences m-passives (hors verbes déponents et passiva tantum) ne possèdent pas de participes présent et futur, ni de gérondif. Une autre différence entre ces classes flexionnelles est la présence (m-actif) ou non (m-passif) d'un impératif de la deuxième personne du pluriel au futur. Ces cas de défectivité systématiques sont indiqués par les cases gris

---

26. Hors cas de déficience ou de supplétion de radical

foncé pour dix paradigmes types dans les tableaux 7.1 à 7.5<sup>27</sup>.

3F Dans 3F, la défektivité concerne seulement les 166 verbes déponents et passiva tantum.

**Périphrase** Les formes réalisées par des procédés périphrastiques :

3H/4H Les paradigmes à formes m-passives sont par ailleurs périphrastiques au perfectif. Nous avons représenté ce fait dans les deux descriptions par une défektivité à un niveau exposant, à savoir le seul niveau d'exponence pour la description 3H et le second niveau d'exponence (le niveau 4 des désinences personnelles) pour la description 4H. Cette défektivité est marquée par les cases gris clair pour cinq paradigmes types dans les tableaux 7.1 à 7.5.

3F En 3F, les cas de périphrase concernent les cases perfectives passives, l'infinitif futur actif et l'infinitif passé passif.

**Synchrétisme** On trouve des cas de synchrétisme dans les deux types d'analyse.

3H/4H Dans les paradigmes verbaux, on observe également un certain nombre de synchrétismes systématiques. Ils concernent systématiquement les impératifs des deuxième et troisième personnes ainsi que les formes des deuxième et troisième personnes du singulier et des première et deuxième personnes du pluriel du futur perfectif de l'indicatif par rapport à ces mêmes personnes au présent perfectif du subjonctif. Notons cependant que pour les verbes ayant un perfectif périphrastique, ce dernier type de synchrétisme n'apparaît pas. Par ailleurs, les verbes à thèmes en *i*, athématiques ou mixtes présentent également un synchrétisme systématique entre la première personne du singulier du présent imperfectif du subjonctif et la case correspondante au futur de l'indicatif. Nous illustrons ces différents synchrétismes pour dix paradigmes type dans les tableaux 7.1 à 7.5 : deux cases de même couleur présentent entre elles un cas de synchrétisme.

3F Dans la mesure où m-actif et m-passif ne constituent en 3F qu'un seul paradigme, les verbes transitifs comportent ainsi jusqu'à sept cas de synchrétisme : 7 pour les verbes transitifs appartenant à la classe flexionnelle  $Z_{v-2}^{exp}$ , six pour ceux de  $Z_{v-1}^{exp}$ .

**Décalage** 3H/4H Nous n'avons pas trouvé de cas de décalage morphosyntaxique en 3H/4H.

---

27. Une autre analyse possible aurait consisté à poser deux catégories flexionnelles distinctes pour les verbes à désinences m-actives et les verbes à désinence m-passives. Dans ce cas, les deux catégories flexionnelles auraient chacune eu leur inventaire de paquets de traits morphosyntaxiques propre. L'opération dérivationnelle de passivation aurait quant à elle provoqué un changement de catégorie flexionnelle.

3F En 3F, en revanche, les verbes déponents présentent une instance systématique de décalage morphosyntaxique. Au niveau exposant, ils sont soumis à une fonction de transfert  $t_{depo}$  qui réalise, par des règles de réalisation du passif, des formes actives.

Notons que le latin présente également un nombre conséquent de cas de surabondance. La distribution de la surabondance est en partie systématique. La marque m-passive de la deuxième personne du singulier *-ris* coexiste ainsi régulièrement avec la marque *-re*. De même, la marque du 3PL.PRS.PFV.IND *-erunt* est également régulièrement réalisée *-ere*. Toutefois, il existe, à côté de ces cas réguliers, un nombre très élevé de cas de surabondance ponctuelle, lexicalement spécifiés. Ces cas demanderaient une étude approfondie de chaque cas particulier qui dépasse largement l'ambition de ce chapitre. Nous avons par conséquent pris le parti de laisser volontairement de côté ici la question de la surabondance, y compris dans sa dimension régulière, plutôt que de proposer des mesures incomplètes de ce phénomène. Les mesures de surabondance seront néanmoins facilement intégrables dans l'inventaire des mesures que nous fournissons ci-dessous dès lors que le phénomène aura été décrit de façon satisfaisante et complète.

Enfin, nous n'avons pas référencé de cas de surdifférenciation <sup>28</sup>.

Pour les cas de non-canonicité décrits ci-dessus, nous avons calculé pour chacun des deux types de description les indices de canonicité associé au système flexionnel verbal du latin. Dans les sections suivantes, nous indiquons les mesures associées au système flexionnel du latin selon chacune de nos deux descriptions. Le tableau 7.14 résume les résultats des mesures de canonicité obtenues pour les deux descriptions. Elles sont également illustrées en bleu pour la description 3H/4H <sup>29</sup> et en vert pour la description 3F dans le diagramme de Kiviat de la figure 7.31.

---

28. Ce fait découle directement de la manière dont nous avons défini les catégories flexionnelles dans les deux analyses. Notamment, nous aurions pu poser une catégorie spécifique pour les verbes intransitifs en 3F qui n'aurait pas comporté de PTM avec la valeur passive, ni de PTM pour le participe passé, l'adjectif ou encore le supin. Dans ce cas, un verbe déponent intransitif aurait présenté de la surdifférenciation pour ces formes non finies.

29. Dans la mesure où les deux analyses 3H et 4H sont identiques en matière de canonicité, nous illustrerons les cas de ces deux analyses à partir d'exemples de 4H. Les exemples de 3H auraient cependant tout autant être pu utilisés pour illustrer les différents cas de non-canonicité.

## 7.4.2 Mesures de la réalisation canonique des traits morphosyntaxiques

### 7.4.2.1 Mesure de la déficience

L'indice de déficience d'un flexème correspond au nombre de paquets de traits morphosyntaxiques (PTM) de l'ensemble des PTM de sa catégorie flexionnelle qui ne font pas également partie de l'ensemble des PTM pour lequel ce flexème se fléchit.

Pour les flexèmes verbaux du latin, l'indice de déficience de chaque flexème dépend de la définition de la catégorie flexionnelle verbale et plus particulièrement de son inventaire de PTM. Les catégories flexionnelles verbales définies pour 3H, et 3F sont respectivement indiquées aux figures 7.1 et 7.25. La catégorie flexionnelle de 4H comporte le même inventaire de PTM que celle de 3H.

**Déficience en 3H/4H** En 3H/4H, la catégorie flexionnelle verbale comporte ainsi un inventaire de 196 PTM si l'on considère l'intégralité des formes des participes, gérondif, adjectif verbal, supin.

COEPI	
I-PHON	coepi
I-CAT	<i>verbe</i>
MSF	$\left\{ \begin{array}{l} (1\vee 2\vee 3).(SG\vee PL).(PRS\vee PST).(IPFV).(IND\vee SBJV), \\ (1\vee 2\vee 3).(SG\vee PL).FUT.(IPFV).SBJV, \\ (PST.INF, (PST\vee FUT).PTCP, \end{array} \right\}$
S-STEM	( <i>vide</i> )
S-FORM	( <i>vide</i> )
I-PAT	$\begin{array}{l} (Z_{REG-EQ-T}^s, id), (Z_c^{th}, id) (Z_{II}^{exp}, id) (Z_A^{exp}, id) (Z_{adj-oa}^{exp}, id) \\ (Z_{REG-EQ-T}^s, id), (Z_c^{th}, id) (Z_{II}^{exp}, id) (Z_A^{exp}, id) (Z_{adj-c}^{exp}, id) \end{array}$

FIGURE 7.30 : Analyse 4H : entrée lexicale du flexème COEPI « avoir commencé »

D'après cette définition de la catégorie flexionnelle verbale, un flexème déficient comme COEPI « avoir commencé », illustré à la figure 7.30 sera déficient pour 68 PTM, en l'occurrence ceux comportant les PTM suivants :

- IPFV.IND, IPFV.SBJV et IMP
- PRS.PTCP, GRV, GRD et SUP



– PRS.INF, FUT.INF

Autrement dit, COEPI possède un inventaire de 128 PTM. Son indice de déficience vaut ainsi :

$$\begin{aligned}\mathcal{D}^-(\text{COEPI}) &= |\Sigma_{\text{VERBE}} \setminus \Sigma_{\text{COEPI}}| \\ &= 196 - 128 \\ &= 68\end{aligned}\tag{7.1}$$

Il n'existe en tout dans *Leffa* que quatre verbes présentant le même type de déficience que COEPI. En revanche, il existe en tout dans nos deux descriptions treize types de déficience différents. Ces treize types sont indiqués au tableau 7.12 avec à chaque fois l'inventaire et le nombre des PTM effectivement compris dans les ensembles de PTM définissant les différents flexèmes déficients et le nombre de flexèmes touchés par ce type de déficience. La déficience de COEPI est étiquetée PFV dans le tableau 7.12.

L'indice de déficience pour l'intégralité d'un système flexionnel est calculé comme la moyenne des indices de déficience de ses flexèmes  $\mathcal{D}^-(i)$  sur le lexique flexionnel. Pour notre système flexionnel verbal du latin en 3H/4H, il vaut :

$$\begin{aligned}\mathcal{D}^-(\mu_{\text{infl}}) &= \frac{\sum_{i \in \text{Leffa}} \mathcal{D}^-(i)}{|\text{Leffa}_{\text{TRAD}}|} \\ &= \frac{4499}{2271} \approx 2,0\end{aligned}\tag{7.2}$$

Il manque ainsi en moyenne 2,0 PTM à chaque paradigme verbal.

**Déficience en 3F** Dans *Leffa*<sub>TRAD</sub>, on trouve en tout 21 types de déficience. Ils sont indiqués au tableau 7.13 avec à chaque fois l'inventaire et le nombre des PTM effectivement compris dans les ensembles de PTM définissant les différents flexèmes déficients et le nombre de flexèmes touchés par ce type de déficience

*Leffa*<sub>TRAD</sub> comporte en tout 1 276 flexèmes. Pour notre système flexionnel verbal du latin en 3F, l'indice de déficience vaut :

30. Les types de déficience marqués comme irrégulier correspondent à des cas où tous les PTM sont stipulés dans le lexique ; il en existe ici de deux types.

31. Les types de déficience marqués comme irrégulier correspondent à des cas où tous les PTM sont stipulés dans le lexique ; il en existe ici de deux types.

TYPE DE DÉFICIENCE	INVENTAIRE DE PTM	NOMBRE DE PTM	NOMBRE DE FLEXÈMES
NO-S3	tout sauf FUT.PTCP, SUP, PST.PTCP	134	15
IPFV	IPFV.IND, IPFV.SBJV, PTCP, GRD, GRV	185	13
NO-S2S3	tout sauf PFV.IND, PFV.SBJV, INF.PST, FUT.PTCP, SUP, PST.PTCP	73	10
IMPERS-SG	3SG.IND, 3SG.SBJV, PTCP, GRD, GRV, INF	138	8
IMPERS-SG-IPFV	3SG.IPFV.IND, 3SG.IPFV.SBJV, PTCP, GRD, GRV	130	4
PFV	PFV.IND, PFV.SBJV, PTCP, IMP, INF sauf PTCP.PRS	128	4
PTCP-PST	PTCP.PST, GRD	60	3
NO-S2	tout sauf PFV.IND, PFV.SBJV, INF.PST	165	2
PTCP	PTCP, GRV, GRD, SUP	127	2
PTCP-PRS-IPRF	PTCP.PRS, PST.IPFV.IND	36	1
IMPERS	3.IND, 3.SBJV, PTCP, GRD, GRV	145	1
NOFEAT	irrégulier <sup>30</sup>	2	1
NOFEAT	irrégulier	13	1

TABLE 7.12 : Types de déficience de la flexion verbale latine selon 3H/4H

$$\begin{aligned}
 \mathcal{D}^-(\mu_{\text{infl}}) &= \frac{\sum_{i \in \text{Leffla}_{\text{TRAD}}} \mathcal{D}^-(i)}{|\text{Leffla}_{\text{TRAD}}|} \\
 &= \frac{16215}{1276} \approx 12,7
 \end{aligned}
 \tag{7.3}$$

Avec en moyenne 12,7 PTM manquants dans les inventaires de PTM des entrées lexicales, 3F est ainsi bien plus déficiente que 3H/4H.

TYPE DE DÉFICIENCE	INVENTAIRE DE PTM	NOMBRE DE PTM	NB DE FLEXÈMES
ACTONLY	ACT, PRS.PTCP.ACT, FUT.PTCP, SUP	196	95
ACTONLY-NO-S3	ACT, PRS.PTCP.ACT, GRV, GRD	134	11
ACTONLY-IPFV	IPFV.IND.ACT, IPFV.SBJV.ACT, PTCP.PRS, PTCP.FUT, SUP, GRV, IMP	103	7
ACTONLY-NO-S2S3	ACT.PFV, INF.ACT, GRV, IMP.ACT	44	7
ACTONLY-IMPERS-SG	3SG.IND.ACT, 3SG.SBJV.ACT, PTCP.PRS, PTCP.FUT, SUP, GRV, INF.ACT	80	7
ACTONLY-IMPERS-SG-IPFV	3SG.IPFV.IND, 3SG.IPFV.SBJV, PTCP.PRS, PTCP.FUT, SUP, GRV, INF.ACT	75	4
IPFV	IPFV.IND, IPFV.SBJV, PTCP, IMP, INF.PRS sauf PTCP.PST	132	3
NO-S2S3	IPFV.IND, IPFV.SBJV, IMP, PTCP.PRS, GRV, GRD, INF.PRS	138	3
NO-S2	IND.IPFV, SBJV.IPFV, IMP, PTCP, SUP, GRV, GRD, INF.PRS, INF.FUT, INF.PST.PASS	201	2
PASSONLY-PTCP-PST	PTCP.PST, GRD	60	2
ACTONLY-PFV	tout ACT sauf IPFV.IND.ACT, IPFV.SBJV.ACT, INF.PST.ACT, et PTCP, GRV, SUP	135	2
ACTONLY-PTCP	PTCP.PRS, PTCP.FUT, GRV, SUP	67	1
ACTONLY-IMPERS	3.IND.ACT, 3.SBJV.ACT	20	1
ACTONLY-IMPERS-SG-AND-PTCP-PST	3SG.IND.ACT, 3SG.SBJV.ACT, PTCP.PST	40	1
PASSONLY-IPFV	IPFV.PASS,PTCP, GRV, GRD, SUP	104	1
ACTONLY-PTCP-PRS	PTCP.PRS	30	1
NO-S3	IND, SBJV, IMP, PTCP.PRS, GRV, GRD, INF.PRS, INF.PST.ACT	201	1
ACTONLY-PTCP-PRS-IPRF	PTCP.PRS, PST.IPFV.IND.ACT	36	1
PASSONLY-PTCP	PTCP.PST, GRD	60	1
PFV	PFV.IND, PFV.SBJV, PTCP.PST, GRV, GRD, SUP, INF.PST	129	1
NOFEAT	irrégulier <sup>31</sup>	2	1
NOFEAT	irrégulier	13	1

TABLE 7.13 : Types de déficience de la flexion verbale latine selon 3F

### 7.4.2.2 Mesure du décalage morphosyntaxique

Les cas de décalage morphosyntaxique sont les cas où pour un flexème donné, les PTM réalisés ne sont pas ceux qui sont canoniquement réalisés par les règles de la zone réalisationnelle employée pour construire les formes. Dans  $\mathbb{P}\mathbb{R}\mathbb{S}\mathbb{L}$ , il s'agit des cas où la fonction de transfert associée à une zone réalisationnelle dans le schème flexionnel d'une entrée n'est pas la fonction identité.

**Décalage morphosyntaxique en 3H/4H** En 3H/4H toutes les fonctions de transfert dans les schèmes flexionnels des entrées sont la fonction identité. L'indice de décalage morphosyntaxique de chacune d'entre elles est donc nul. L'indice de décalage morphosyntaxique du système complet est également nul.

$$\begin{aligned} \mathcal{M}(\mu_{\text{infl}}) &= \frac{\sum_{i \in \text{Le}fa} \mathcal{M}(i)}{|\text{Le}fa|} \\ &= 0 \end{aligned} \quad (7.4)$$

**Décalage morphosyntaxique en 3F** Dans 3F, les verbes déponents comme LOQUOR « parler » possèdent dans leur schème flexionnel un couple réalisationnel de niveau exponence pour lequel la fonction de transfert n'est pas la fonction identité mais la fonction  $t_{\text{depo}}$  définie comme suit :

$$t_{\text{depo}}(\text{ACT}) = \text{PASS} \quad (7.5)$$

L'indice de décalage morphosyntaxique d'un flexème  $i$  donné est obtenu par le nombre de partitions de l'ensemble  $\Sigma_{e_i}$  tel que la fonction de transfert associée à chacune de ces partitions n'est pas la fonction identité à un niveau donné au moins.

Celui du verbe comme LOQUOR vaut alors :

$$\begin{aligned} \mathcal{M}(\text{LOQUOR}) &= |\{\Sigma \in \mathcal{P}_{\Sigma_{\text{LOQUOR}}} \mid (\exists l_{\Sigma} \in \{l_{e_{\text{VERBE}},1}, \dots, l_{e_{\text{VERBE}},n}\} \mid \\ &\quad \forall \sigma \in \Sigma, (z_{\sigma, l_{\Sigma}}, t_{\sigma, l_{\Sigma}}) \in \mathbb{L}_{l_{\Sigma}}^{\text{LOQUOR}}(\Sigma_{\text{LOQUOR}}) \Rightarrow t_{\sigma, l_{\Sigma}} \neq \text{id})\}| \\ &= 1 \end{aligned} \quad (7.6)$$

Il existe dans  $\text{Le}ffa_{\text{TRAD}}$  166 flexèmes de ce type. L'indice de décalage morphosyntaxique

du système flexionnel complet dans 3F vaut ainsi :

$$\begin{aligned} \mathcal{M}(\mu_{\text{infl}}) &= \frac{\sum_{i \in \text{Leff}a_{\text{TRAD}}} \mathcal{M}(i)}{|\text{Leff}a_{\text{TRAD}}|} \\ &= \frac{166}{1276} \approx 0,13 \end{aligned} \quad (7.7)$$

### 7.4.3 Mesures du remplissage des cases des paradigmes

#### 7.4.3.1 Mesure de la défektivité

La défektivité se calcule par rapport à l'intégralité des PTM effectivement exprimés par un flexème donné. Pour chaque flexème  $i$ , l'indice de défektivité  $\mathcal{F}^-(i)$  de  $i$  est calculé comme le nombre de PTM qui ne sont pas associés à un sous-schème complet dans son schème.

$$\mathcal{F}^-(i) = \frac{|\{\sigma \in \Sigma_{e_i} \cap \Sigma_i \mid \exists l \in \{l_1, \dots, l_n\}^{e_i} \mid \mathbb{L}_l^i(\sigma) = \emptyset\}|}{|\Sigma_{e_i}|} \quad (7.8)$$

**Défektivité en 3H/4H** En 3H/4H, la défektivité est un phénomène systématique qui touche l'ensemble des flexèmes non hétéroclites, autrement dit, qui ne sont ni des verbes déponents ni des passiva tantum.

Les 1 065 verbes à désinences m-actives sont ainsi défectifs pour les PTM suivants :

- PST.PTCP, GRD et FUT.INF

Cela revient ainsi à 61 PTM non associés à une zone flexionnelle par verbe.

Les 976 verbes à désinences m-passives sont quant à eux défectifs pour les PTM suivants :

- PRS.PTCP et FUT.PTCP
- PFV.IND et PFV.SBJV, FUT.INF, PST.INF, GRV, et SUP
- 2PL.FUT.IPFV.IMP

Cela revient ainsi à 100 PTM non associés à une zone flexionnelle par verbe.

Les deux ensembles de verbes non hétéroclites ont ainsi des cases vides pour les cases associées à ces PTM. Les autres cas des cases vides dans les paradigmes ont été analysés comme des cas de déficience décrits à la section précédente.

Les verbes déponents et les passiva tantum ne sont pas défectifs. Leur indice de défektivité est donc nul.

Les verbes à désinences m-actives non défectifs sont défectifs pour 61 paquets de traits morphosyntaxiques. Or, l'indice de défectivité formelle d'un paradigme se calcule comme le nombre de cases défectives divisé par le nombre canonique de cases du paradigme. À l'exemple de l'indice de défectivité du flexème AMO, nous pouvons ainsi calculer l'indice de défectivité de chacun d'entre eux :

$$\begin{aligned}\mathcal{F}^-(\text{AMO}) &= \frac{|\{\sigma \in \Sigma_{\text{VERBE}} \cap \Sigma_{\text{AMOR}} \mid \exists l \in \{l_{\text{VERBE},1}, \dots, l_{\text{VERBE},n}\} \mid \mathbb{L}_l^{\text{AMO}}(\sigma) = \emptyset\}|}{|\Sigma_{\text{VERBE}}|} \\ &= \frac{61}{196} \approx 31,1\%\end{aligned}\tag{7.9}$$

L'indice de défectivité de AMOR vaut quant à lui :

$$\begin{aligned}\mathcal{F}^-(\text{AMOR}) &= \frac{|\{\sigma \in \Sigma_{\text{VERBE}} \cap \Sigma_{\text{AMOR}} \mid \exists l \in \{l_{\text{VERBE},1}, \dots, l_{\text{VERBE},n}\} \mid \mathbb{L}_l^{\text{AMOR}}(\sigma) = \emptyset\}|}{|\Sigma_{\text{VERBE}}|} \\ &= \frac{100}{263} = 38,0\%\end{aligned}\tag{7.10}$$

Nous obtenons alors l'indice de défectivité globale de notre système flexionnel en calculant la moyenne des indices de défectivité associés à chacun de ses flexèmes  $i$  sur notre lexique flexionnel  $\text{Leffla}$ .

$$\begin{aligned}\mathcal{F}^-(\mu_{\text{infl}}) &= \frac{\sum_{i \in \text{Leffa}} \mathcal{F}^-(i)}{|\text{Leffa}|} \\ &= \frac{1065 \times \frac{61}{196} + 976 \times \frac{100}{263}}{2271} \approx 34,3\%\end{aligned}\tag{7.11}$$

**Défectivité en 3F** En 3H, seuls les 166 verbes déponents et les passiva tantum sont également défectifs. Les autres cases manquantes sont considérées comme des cas de déficience. À chacun des flexèmes défectifs de 3F, il manque les 66 formes m-actives (dont les six formes impératives). L'indice de défectivité de 3F vaut ainsi :

$$\mathcal{F}^-(\mu_{\text{infl}}) = \frac{\sum_{i \in \text{Leff}a_{\text{TRAD}}} \mathcal{F}^-(i)}{|\text{Leff}a_{\text{TRAD}}|} \quad (7.12)$$

$$\frac{166 \times \frac{66}{263}}{1276} \approx 3,3\%$$

On constate ainsi que, contrairement à 3H/4H, 3F est quasiment non défective. Mais ce fait provient en très grande partie du fait que les verbes transitifs et intransitifs réguliers ont été traités comme déficients et non comme défectifs. Nous avons fait ce choix à cause de l'existence des deux classes flexionnelles complètes de 3F qui comprennent à chaque fois toutes les formes non finies, contrairement aux classes de 3H/4H qui ne comprenaient à chaque fois de règles de réalisation pour une partie des formes non finies seulement.

#### 7.4.3.2 Mesure de l'allomorphie radicale

L'indice d'allomorphie régulière du paradigme d'un flexème donné est nul lorsque ce flexème n'affiche dans son paradigme qu'un seul radical ou lorsque les radicaux différents sont des radicaux supplétifs.

Pour un flexème  $i$  donné, l'indice d'allomorphie régulière se calcule ainsi selon la formule suivante :

$$\mathcal{A}(i) = |\mathcal{A}_{\text{set}}(i)| - 1 \quad (7.13)$$

où  $\mathcal{A}_{\text{set}}(i)$  est l'ensemble des radicaux réguliers appartenant au flexème  $i$ .

**Allomorphie radicale en 3H/4H** Dans nos descriptions, l'indice d'allomorphie régulière sera nul pour tous les flexèmes appartenant soit à la classe de radicaux  $Z_{\text{REG-EQ-EQ}}^s$  (c'est-à-dire qui ne manifestent aucune alternance radicale) soit à la classe de radicaux  $Z_{\text{SUPPL}}^s$  entièrement supplétives.

Pour un verbe comme DUCO « mener, conduire » qui possède dans son paradigme trois radicaux (*duc-*, *dux* et *duct*) réguliers construits selon la classe de radicaux  $Z_{\text{REG-X-T}}^s$ , l'indice d'allomorphie régulière vaut :

$$\begin{aligned} \mathcal{A}(\text{DUCO}) &= |\mathcal{A}_{\text{set}}(\text{DUCO})| - 1 \\ &= 3 - 1 = 2 \end{aligned} \quad (7.14)$$

Selon leurs classes de radicaux et si l'on néglige pour le moment les cas de supplétion de radical, l'indice d'allomorphie régulière des différents flexèmes vaut ainsi :

- 2 pour les verbes réguliers comme DUCO présentant chacun trois radicaux réguliers distincts ; ces verbes sont au nombre de 1 664.
- 1 pour les verbes dont la classe de radicaux comporte une zone de radical différente de celle construisant le RAD1 commune avec la classe de radicaux  $Z_{\text{REG-EQ-EQ}}^s$  ; un exemple de cela est le verbe NECTO « tisser » appartenant à la classe de radicaux  $Z_{\text{REG-X-EQ}}^s$  avec pour radicaux réguliers *nect-*, *nex-* et *nect-* ; il y a de 395 verbes de ce type.
- 1 pour les verbes ayant deux radicaux identiques pour le RAD2 et le RAD3 comme FLECTO « tordre » dont les radicaux sont *flect-*, *flex* et *flex* ; ces verbes sont au nombre de 78.
- 0 pour les verbes appartenant à la classe de radicaux  $Z_{\text{REG-EQ-EQ}}^s$  comme VISO « regarder » ou à la classe  $Z_{\text{SUPPL}}^s$  comme FERRO « porter » dont les radicaux supplétifs sont respectivement *tul-* pour le RAD2 et *-lat* pour le RAD3 ; il y en a 134.

L'indice d'allomorphie radicale régulière d'un système flexionnel est la moyenne des indices d'allomorphie radicale des flexèmes de son lexique flexionnel. Pour tout flexème  $i$  appartenant à un lexique flexionnel  $\mathcal{L}$ , on le calcule suivant la formule ci-dessous :

$$\mathcal{A}(\mu_{\text{infl}}) = \frac{\sum_{i \in \mathcal{L}} \mathcal{A}(i)}{|\mathcal{L}|} \quad (7.15)$$

Au chapitre 4, nous avons par ailleurs défini l'ensemble des radicaux allomorphiques réguliers pour un flexème donné comme la différence entre le nombre de ses radicaux distincts et le nombre de ses radicaux supplétifs.

$$S_{\text{dist}}(i) = S_{\text{dist}}^{\text{reg}}(i) \cup S_i^{\text{suppl}} \quad (7.16)$$

Sachant que par définition

$$S_{\text{dist}}^{\text{reg}}(i) \cap S_i^{\text{suppl}} = \emptyset \quad (7.17)$$

Pour un système flexionnel donné, nous pouvons ainsi calculer l'indice d'allomorphie régulière comme la moyenne des indices de l'ensemble des flexèmes sur la taille du lexique



flexionnel :

$$\mathcal{A}(\mu_{\text{infl}}) = \frac{\sum_{i \in \mathcal{I}} \mathcal{A}(i)}{|\mathcal{I}|} \quad (7.18)$$

Dans le cas de notre lexique flexionnel  $\text{Leffla}$ , l'indice d'allomorphie régulière du système verbal latin vaut :

$$\begin{aligned} \mathcal{A}(\mu_{\text{infl}}) &= \frac{\sum_{i \in \text{Leffa}} \mathcal{A}(i)}{|\text{Leffa}|} \\ &= \frac{2 \times 1664 + 1 \times 395 + 1 \times 75 + 0 \times 134}{2271} \approx 1,67 \end{aligned} \quad (7.19)$$

**Allomorphie radicale en 3F** Nous pouvons également calculer l'indice d'allomorphie régulière comme la différence entre la taille de l'ensemble des radicaux distincts et l'ensemble des supplétifs pour tous les flexèmes moins le nombre de flexèmes du lexique flexionnel divisé par la taille du lexique flexionnel.

$$\begin{aligned} \mathcal{A}(\mu_{\text{infl}}) &= \frac{\sum_{i \in \mathcal{I}} \mathcal{A}(i)}{|\mathcal{I}|} \\ &= \frac{\sum_{i \in \mathcal{I}} |S_{\text{dist}}(i)| - \sum_{i \in \mathcal{I}} |S_i^{\text{suppl}}| - |\mathcal{I}|}{|\mathcal{I}|} \end{aligned} \quad (7.20)$$

Dans le cas de  $\text{Leffla}_{\text{TRAD}}$ , l'indice d'allomorphie régulière du système verbal latin vaut alors :

$$\begin{aligned} \mathcal{A}(\mu_{\text{infl}}) &= \frac{\sum_{i \in \text{Leffla}_{\text{TRAD}}} \mathcal{A}(i)}{|\text{Leffla}_{\text{TRAD}}|} \\ &= \frac{\sum_{i \in \text{Leffla}_{\text{TRAD}}} |S_{\text{dist}}(i)| - \sum_{i \in \text{Leffla}_{\text{TRAD}}} |S_i^{\text{suppl}}| - |\text{Leffa}|}{|\text{Leffla}_{\text{TRAD}}|} \\ &= \frac{4544 - 229 - 1276}{1276} = \frac{3039}{1276} \approx 2,38 \end{aligned} \quad (7.21)$$

La différence entre les indices d'allomorphie des deux analyses provient essentiellement du plus petit nombre d'entrées lexicales dans 3F par rapport à 3H/4H. Le poids des flexèmes irréguliers sur le lexique flexionnel dans son ensemble est ainsi bien plus important dans 3F.

### 7.4.3.3 Mesure de l'hétéroclise

L'hétéroclise peut se mesurer niveau par niveau pour chaque paradigme flexionnel donné. L'indice d'hétérocliticité d'un niveau réalisationnel  $l$  pour un flexème  $i$  donné est défini comme le nombre de classes flexionnelles auxquelles appartiennent les zones réalisationnelles pour lesquelles  $i$  se fléchit.

**Hétéroclise en 3H/4H** Dans 3H/4H, les flexèmes peuvent être hétéroclites au niveau exposant lorsqu'ils sont des verbes déponents ou des passiva tantum. Dans l'analyse 3H, les flexèmes hétéroclites se fléchissent au niveau 3 soit selon des zones flexionnelles appartenant aux classes d'exponence  $Z_{v-A}^{exp}$  et  $Z_{v-B}^{exp}$  soit selon des zones appartenant aux classes d'exponence  $Z_{v-C}^{exp}$  et  $Z_{v-D}^{exp}$ . Dans l'analyse 4H, tous les flexèmes hétéroclites au niveau exposant se fléchissent au niveau 4 selon les classes d'exponence  $Z_A^{exp}$  et  $Z_B^{exp}$ . Un flexème comme OPINOR « *penser* » aura ainsi au niveau exposant (niveau 3 en 3H et niveau 4 en 4H) un indice d'hétérocliticité calculé selon la formule suivante <sup>32</sup> :

$$\begin{aligned} \mathcal{H}_{niv\ 4}(\text{OPINOR}) &= |\{Z_A^{exp}, Z_B^{exp}\}| - 1 \\ &= 2 - 1 = 1 \end{aligned} \tag{7.22}$$

Il y a en tout 166 flexèmes qui sont ainsi hétéroclites au niveau d'exponence.

De façon orthogonale à cela, les flexèmes à thème mixtes sont par ailleurs hétéroclites au niveau thématique. Un flexème comme ACCIPIO « *accepter* » a ainsi l'indice d'hétérocliticité suivant :

$$\begin{aligned} \mathcal{H}_{niv\ 2}(\text{ACCIPIO}) &= |\{Z_i^{th}, Z_c^{th}\}| - 1 \\ &= 2 - 1 = 1 \end{aligned} \tag{7.23}$$

Dans notre lexique flexionnel *Leffa* 126 flexèmes sont ainsi hétéroclites au niveau thématique.

L'indice d'hétérocliticité global  $\mathcal{H}(i)$  d'un flexème  $i$  est ensuite défini comme la somme de ses indices d'hétérocliticité par niveau  $\mathcal{H}_l(i)$  pour l'ensemble des niveaux réalisationnels intervenant dans sa flexion.

Un flexème qui serait à la fois hétéroclite au niveau thématique et exposant aurait alors un indice d'hétérocliticité de  $2 + 2 = 4$ . Toutefois, il n'existe pas, dans *Leffa*, de flexème qui

32. Nous utilisons ici le nom des classes du niveau 4 de l'analyse 4H.

combine ces deux types d'hétéroclise.

Nous pouvons enfin calculer l'indice d'hétéroclivité global  $\mathcal{H}(\mu_{\text{infl}})$  de notre système flexionnel verbal du latin. Il correspond à la moyenne des indices d'hétéroclivité de ses flexèmes  $i$  sur son lexique flexionnel  $\text{Leffla}$ .

$$\begin{aligned}\mathcal{H}(\mu_{\text{infl}}) &= \frac{\sum_{i \in \text{Leffa}} \mathcal{H}(i)}{|\text{Leffa}|} \\ &= \frac{166 \times 1 + 126 \times 1}{2271} \approx 0,13\end{aligned}\tag{7.24}$$

**Hétéroclise en 3F** En 3F, il n'y a pas d'instances d'hétéroclise. L'indice d'hétéroclivité de 3F est donc nul.

$$\begin{aligned}\mathcal{H}(\mu_{\text{infl}}) &= \frac{\sum_{i \in \text{Leffla}_{\text{TRAD}}} \mathcal{H}(i)}{|\text{Leffla}_{\text{TRAD}}|} \\ &= 0\end{aligned}\tag{7.25}$$

## 7.4.4 Mesures de la canonicité des réalisations

### 7.4.4.1 Mesure des cas de supplétion de radicaux

L'indice de supplétion de radicaux d'un flexème donné correspond au nombre de radicaux supplétifs dans son entrée lexicale. Cet indice est donc calculé directement sur le lexique flexionnel.

**Supplétion de radicaux en 3H/4H** Dans l'analyse 3H/4H, pour le verbe `POSSUM` « pouvoir », l'ensemble des radicaux supplétifs indiqués dans son entrée comporte trois radicaux : *potu-*, *possu-* et *pote-*. Son indice de supplétion vaut ainsi :

$$\mathcal{R}^s(\text{POSSUM}) = |\mathcal{S}_{\text{POSSUM}}^{\text{suppl}}| = 3\tag{7.26}$$

L'indice de supplétion de radicaux d'un système flexionnel complet correspond à la moyenne des indices de supplétion de radicaux de ses flexèmes sur son lexique flexionnel. En d'autres termes, l'indice de supplétion de radicaux d'un système flexionnel donné est la somme des radicaux listés dans le lexique flexionnel divisé par le nombre de flexèmes.

*Leffa* spécifie en tout 377 radicaux. L'indice de supplétion de radicaux dans 3H/4H vaut ainsi :

$$\begin{aligned} \mathcal{R}^s(\mu_{\text{infl}}) &= \frac{\sum_{i \in \text{Leffa}} \mathcal{R}^s(i)}{|\text{Leffa}|} \\ &= \frac{377}{2271} \approx 0,3 \end{aligned} \quad (7.27)$$

Supplétion de radicaux en 3F *Leffa*<sub>TRAD</sub> spécifie en tout 229 radicaux. L'indice de supplétion de radicaux dans 3F vaut donc :

$$\begin{aligned} \mathcal{R}^s(\mu_{\text{infl}}) &= \frac{\sum_{i \in \text{Leffa}} \mathcal{R}^s(i)}{|\text{Leffa}|} \\ &= \frac{229}{1276} \approx 0,18 \end{aligned} \quad (7.28)$$

Cette différence assez importante peut paraître surprenante compte tenu de la faible différence de compacité pour la composante supplétion entre 3H/4H d'une part et 3F de l'autre. Ce faible impact s'explique par le poids quasi-négligeable de la supplétion de radicaux par rapport à la supplétion de formes dans la longueur de description des analyses. Or, comme nous le l'indiquons ci-dessous, les lexiques flexionnels des deux types d'analyses comportent le même nombre de formes supplétives<sup>33</sup>.

#### 7.4.4.2 Mesure des cas de supplétion de forme

Supplétion de forme en 3H/4H Pour un flexème donné, l'indice de supplétion de forme s'obtient à partir du nombre de formes spécifiées dans son entrée lexicale.

Pour le verbe POSSUM « *pouvoir* », le nombre de formes supplétives spécifiées est de 27. Son indice de supplétion de formes vaut ainsi :

$$\mathcal{R}^f(\text{POSSUM}) = |F_{\text{POSSUM}}^{\text{suppl}}| = 27 \quad (7.29)$$

L'indice de supplétion de forme pour l'intégralité du lexique flexionnel est quant à lui calculé à partir de l'ensemble des formes supplétives qu'il contient. *Leffa* comporte 410 formes supplétives. Il vaut :

33. Les formes supplétives dans les deux lexiques flexionnels sont des formes m-actives.

$$\mathcal{R}^f(\mu_{\text{infl}}) = \frac{\sum_{i \in \text{Leffa}} \mathcal{R}^f(i)}{|\text{Leffa}|} \quad (7.30)$$

$$\frac{410}{2271} \approx 0,18$$

Supplétion de forme en 3F  $\text{Leffa}_{\text{TRAD}}$  comporte également 410 formes supplétives. Son indice de supplétion de forme vaut donc :

$$\mathcal{R}^f(\mu_{\text{infl}}) = \frac{\sum_{i \in \text{Leffa}} \mathcal{R}^f(i)}{|\text{Leffa}|} \quad (7.31)$$

$$\frac{410}{1276} \approx 0,32$$

Compte tenu des tailles respectives de  $\text{Leffa}$  et  $\text{Leffa}_{\text{TRAD}}$ , 3F obtient en raison de son lexique flexionnel plus petit un indice de supplétion de forme plus important que 3H/4H.

#### 7.4.4.3 Mesure des syncrétismes

Des cases syncrétiques sont des cases qui partagent pour deux PTM distincts une même réalisation phonique. Pour un flexème donné son indice de syncrétisme  $\mathcal{S}(i)$  se déduit de la somme des tailles des ensembles syncrétiques  $\mathcal{S}_F(i, r^f)$  de ses cases, chaque ensemble de cases syncrétiques comptant pour sa taille moins 1 dans l'indice.

Syncrétisme en 3H/4H Comme le montrent les tableaux 7.1 à 7.4 plus haut, les flexèmes à désinences m-actives et thème en  $-a$  ou  $-e$  comme AMO « *aimer* » et CONTINEO « *contenir* » présentent six ensembles de cases syncrétiques. Ces cases sont les suivantes :

- 1SG.FUT.IPFV.IND et 1SG.PRS.IPFV.SBJV,
- 2SG.FUT.PFV.IND et 2SG.PRS.PFV.SBJV,
- 3SG.FUT.PFV.IND et 3SG.PRS.PFV.SBJV,
- 1PL.FUT.PFV.IND et 1PL.PRS.PFV.SBJV,
- 2PL.FUT.PFV.IND et 2PL.PRS.PFV.SBJV,
- 2SG.FUT.IMP et 3SG.FUT.IMP,
- 2PL.FUT.IMP et 3PL.FUT.IMP,

Les verbes à désinences m-actives mais à thème en *-i* ou athématiques comme AUDIO « *entendre* », DUCO « *mener, conduire* » ou encore le verbe mixte ACCIPIO « *accepter* » présentent même sept ensembles de cases syncrétiques, à savoir :

- 1SG.FUT.IPFV.IND et 1SG.PRS.IPFV.SBJV,
- 2SG.FUT.PFV.IND et 2SG.PRS.PFV.SBJV,
- 3SG.FUT.PFV.IND et 3SG.PRS.PFV.SBJV,
- 1PL.FUT.PFV.IND et 1PL.PRS.PFV.SBJV,
- 2PL.FUT.PFV.IND et 2PL.PRS.PFV.SBJV,
- 2SG.FUT.IMP et 3SG.FUT.IMP,
- 2PL.FUT.IMP et 3PL.FUT.IMP,

Pour les verbes à désinences m-passives, le nombre d'ensemble des cases syncrétiques est plus réduit. Il y en a deux pour les verbes à thème en *-i* ou athématiques comme AUDIOR « *être entendu* », DUCOR « *être mené, conduit* » ou encore le verbe mixte ACCIPIOR « *être accepté* » :

- 1SG.FUT.IPFV.IND et 1SG.PRS.IPFV.SBJV,
- 2SG.FUT.IMP et 3SG.FUT.IMP,

Il n'y en a plus qu'un pour les verbes à thème en *-a* ou *-e* comme AMOR « *être aimé* » et CONTINEOR « *être contenu* » :

- 2SG.FUT.IMP et 3SG.FUT.IMP,

Dans tous ces ensembles, le nombre de cases syncrétiques est à chaque fois de 2.

Par ailleurs, les participes, gérondif et adjectif verbal empruntent tous les classes flexionnelles adjectivales  $Z_{adj-oa}^{exp}$  et  $Z_{adj-c}^{exp}$ . Comme les adjectifs, ils sont donc hautement syncrétiques.

Le participe présent est syncrétique pour les formes suivantes :

- ACC.SG.M, NOM.SG.N et ACC.SG.N,
- GEN.SG.M, NOM.PL.M et GEN.PL.N,
- DAT.SG.M, ABL.SG.M, DAT.SG.N et ABL.SG.N,
- GEN.PL.M et GEN.PL.N,
- DAT.PL.M, ABL.PL.M, DAT.PL.F, ABL.PL.F, DAT.PL.N et ABL.PL.N,
- NOM.SG.F, ABL.SG.F, NOM.PL.N et ACC.PL.N,

- GEN.SG.F, DAT.SG.F et NOM.PL.F,

Les participes passé et futur et l'adjectif verbal sont syncrétiques pour les formes suivantes :

- NOM.SG.M, NOM.SG.F, NOM.SG.N et ACC.SG.N,
- ACC.SG.M et ACC.SG.F,
- GEN.SG.M, GEN.SG.F et GEN.SG.N,
- DAT.SG.M, DAT.SG.F, DAT.SG.N, ABL.SG.M, ABL.SG.F et ABL.SG.N,
- GEN.PL.M et GEN.PL.N,
- NOM.PL.M, NOM.PL.F, ACC.PL.M et ACC.PL.F,
- NOM.PL.N et ACC.PL.N,
- GEN.PL.M, GEN.PL.F et GEN.PL.N,
- DAT.PL.M, DAT.PL.F, DAT.PL.N, ABL.PL.M, ABL.PL.F et ABL.PL.N,

Le gérondif est syncrétique pour les formes suivantes :

- DAT.SG.M, et ABL.SG.M,

À partir de ces ensembles syncrétiques, nous pouvons calculer les indices de syncrétisme des différents flexèmes.

Ainsi les indices de syncrétisme  $\mathcal{S}(\text{AUDIO})$ ,  $\mathcal{S}(\text{AMO})$ ,  $\mathcal{S}(\text{AUDIOR})$  et  $\mathcal{S}(\text{AMOR})$  respectivement pour les quatre flexèmes AUDIO, AMO, AUDIOR et AMOR valent respectivement :

$$\begin{aligned}\mathcal{S}(\text{AUDIO}) &= \sum_{r^f \in F_{dist}(\text{AUDIO})} \left( |\mathcal{S}_F(\text{AUDIO}, r^f)| - 1 \right) \\ &= 44\end{aligned}\tag{7.32}$$

$$\begin{aligned}\mathcal{S}(\text{AMO}) &= \sum_{r^f \in F_{dist}(\text{AMO})} \left( |\mathcal{S}_F(\text{AMO}, r^f)| - 1 \right) \\ &= 43\end{aligned}\tag{7.33}$$

$$\begin{aligned}\mathcal{S}(\text{AUDIOR}) &= \sum_{r^f \in F_{dist}(\text{AUDIOR})} \left( |\mathcal{S}_F(\text{AUDIOR}, r^f)| - 1 \right) \\ &= 38\end{aligned}\tag{7.34}$$

$$\begin{aligned} \mathcal{S}(\text{AMOR}) &= \sum_{r^f \in F_{\text{dist}}(\text{AMOR})} \left( |\mathcal{S}_F(\text{AMOR}, r^f)| - 1 \right) \\ &= 37 \end{aligned} \quad (7.35)$$

Une autre façon de calculer l'indice de synchrétisme d'un flexème est de faire la différence entre le nombre total de cases et le nombre de cases à réalisation distincte. Pour AUDIO, le nombre total de cases non vides est de 135. Le nombre de réalisations distinctes dans ces cases vaut 91. L'indice de synchrétisme d'AUDIO vaut ainsi :

$$\mathcal{S}(\text{AUDIO}) = 135 - 91 = 44 \quad (7.36)$$

Nous pouvons par cette même méthode calculer l'indice de synchrétisme  $\mathcal{S}(\mu_{\text{infl}})$  de notre système flexionnel verbal en 3H/4H. Nous savons que l'indice de synchrétisme pour un système flexionnel  $\mu_{\text{infl}}$  complet est défini comme la moyenne de l'indice de synchrétisme  $\mathcal{S}(i)$  de tous les lexèmes sur l'ensemble du lexique flexionnel. Or, *Leffa* produit en tout 263 677 formes fléchies dont 159 256 formes distinctes. La différence entre ces deux chiffres correspond à la somme des indices de synchrétisme de l'ensemble des flexèmes de *Leffa*.

L'indice de synchrétisme de notre système flexionnel verbal vaut ainsi :

$$\begin{aligned} \mathcal{S}(\mu_{\text{infl}}) &= \frac{\sum_{i \in \text{Leffa}} \mathcal{S}(i)}{|\text{Leffa}|} \\ &= \frac{263677 - 159256}{2271} \approx 39,6 \end{aligned} \quad (7.37)$$

Synchrétisme en 3F – *Leffa* et *Leffa*<sub>TRAD</sub> produisent le même inventaire de formes. Pour la description 3F, l'indice de synchrétisme  $\mathcal{S}(i)$  vaut ainsi :

$$\begin{aligned} \mathcal{S}(\mu_{\text{infl}}) &= \frac{\sum_{i \in \text{Leffa}_{\text{TRAD}}} \mathcal{S}(i)}{|\text{Leffa}_{\text{TRAD}}|} \\ &= \frac{263677 - 128012}{1276} \approx 106 \end{aligned} \quad (7.38)$$



## 7.4.4.4 Récapitulatif des mesures de canonicité obtenues

Nous avons calculé l'ensemble des mesures de canonicité de sept phénomènes non canoniques pour l'ensemble de notre système flexionnel. Nous résumons les résultats de ces mesures au tableau 7.14. Les résultats sont également illustrés par le diagramme de Kiviat de la figure 7.31. Dans le diagramme, les mesures sur les axes ont été normalisées pour que la valeur maximale des deux mesures soit égale à 1 pour en faciliter la lecture.

PHÉNOMÈNE	INDICE	3H/4H	3F
DÉFICIENCE	$\mathcal{D}^-(\mu_{\text{infl}})$	2,0	12,7
DÉCALAGE MORPHOSYNTAXIQUE	$\mathcal{M}(\mu_{\text{infl}})$	0	0,13
DÉFECTIVITÉ	$\mathcal{F}^-(\mu_{\text{infl}})$	34,3%	3,3%
ALLOMORPHIE RADICALE	$\mathcal{A}(\mu_{\text{infl}})$	1,67	2,38
HÉTÉROCLISE	$\mathcal{H}(\mu_{\text{infl}})$	0,13	0
SUPPL. DE RADICAL	$\mathcal{R}^s(\mu_{\text{infl}})$	0,3	0,18
SUPPL. DE FORME	$\mathcal{R}^f(\mu_{\text{infl}})$	0,18	0,32
SYNCRÉTISME	$\mathcal{S}(\mu_{\text{infl}})$	39,6	106

TABLE 7.14 : Récapitulatif des mesures de canonicité pour le système flexionnel verbal latin

Les chiffres que nous obtenons pour les mesures de canonicité de nos deux types de descriptions montrent que selon la description que nous donnons d'un système flexionnel particulier, les proportions de phénomènes non canoniques peuvent grandement varier.

Parmi les exemples les plus marquants figure ici la distribution quasi-complémentaire de la déficience et de la défektivité dans nos deux analyses. Le fait que ce soit ces deux phénomènes qui, entre autres, puissent présenter une telle distribution n'est cependant pas pour nous surprendre, dans la mesure où tous deux ont trait à l'existence de cases vides dans le paradigme. Cependant, l'une d'entre elles prévoit l'origine de ce vide dans l'inventaire de traits à réaliser là où dans l'autre il est la conséquence d'une absence de zones réalisationnelles, et donc de règles réalisationnelles, associées.

Un autre exemple est celui de l'hétéroclise et du décalage qui n'existent à chaque fois que dans l'un des deux types de descriptions.

Par ailleurs, nous constatons clairement que la taille du lexique flexionnel influe sur les mesures de canonicité, ces dernières étant définies pour les systèmes flexionnels en tant que moyennes par rapport au nombre de flexèmes : deux lexiques flexionnels comme  $Leffla$  et  $Leffla_{\text{TRAD}}$  qui comportent le même nombre de radicaux et de formes supplétives ne produiront ainsi pas les mêmes indices de supplétion en raison de leur différence

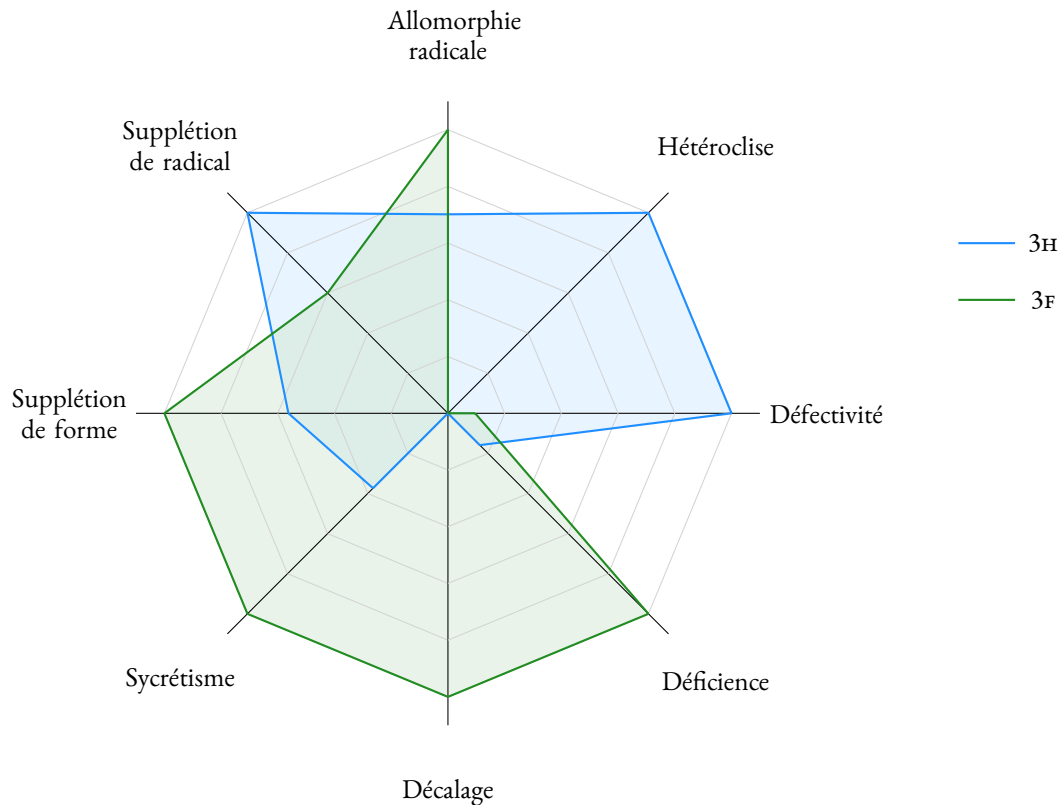


FIGURE 7.31 : Mesures de canonicité de nos descriptions de la flexion verbale latine : Pour faciliter la comparaison, chaque axe est gradué de telle sorte que le maximum des deux valeurs correspondantes pour les deux descriptions fasse 1.

importante en nombre d'entrées lexicales. Ces mesures de canonicité sont ainsi hautement sensibles à la taille du lexique flexionnel.

Nous constatons cependant qu'à ce stade nous ne pouvons tirer aucune conclusion sur la canonicité du système flexionnel verbal du latin à partir de ces deux descriptions, et ce d'autant plus que, comme nous l'avons indiqué en conclusion de la section précédente, les données modélisées ne sont pas les mêmes d'une description à l'autre. 3H/4H a été conçue dans le souci explicite de différencier flexion et dérivation. L'analyse 3F ne fait aucun effort à cet égard. Les données étant ainsi différentes, l'écart observé entre les mesures de canonicité des deux types de descriptions n'aura par conséquent qu'une portée très limitée.

## 7.5 Discussion et perspectives

Dans ce chapitre nous avons comparé trois descriptions concurrentes de la flexion verbale latine. Deux d'entre elles s'appuient sur l'analyse dérivationnelle du passif en latin que nous avons posée au chapitre 1. La première des deux analyses, 3H, a été critiquée pour ce qu'elle présente de la redondance dans le classement des entrées lexicales, réparties d'une part selon leur classe thématique et d'autre part selon leur appartenance à quatre classes d'exponence. Cette redondance se manifeste notamment dans l'association systématique entre les classes thématiques et les classes d'exponence. Par ailleurs, deux de ces quatre classes ont la propriété d'exprimer le marquage m-actif et deux le marquage m-passif, et ce, alors même que la différence ne concerne que quelques rares formes. Enfin, cette critique est d'autant plus marquée que la dérivation passive maintient un clivage entre les deux paires de classes d'exponence en fonction de l'appartenance thématique, et ce quand bien même la dérivation passive était associée à un changement de classe flexionnelle. Ce changement est rappelé schématiquement à la figure 7.32.

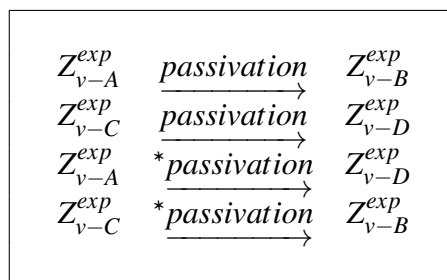


FIGURE 7.32 : Attribution des classes flexionnelles pour les flexèmes des verbes m-actifs dérivés dans l'analyse 3H

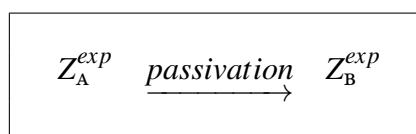


FIGURE 7.33 : Attribution des classes flexionnelles pour les flexèmes des verbes m-actifs dérivés dans l'analyse 4H

En réponse à cette critique, nous avons construit une description concurrente de la flexion verbale latine, la description 4H. Cette description se distingue de la première en ce qu'elle répartit les exposants sur deux niveaux réalisationnels dont seul le second (le niveau 4) exprime les marques de personne et de nombre. Du point de vue de la représentation de la dérivation passive, cette analyse semble à première vue plus

simple et moins redondante. Sa réalisation formelle correspond à un unique changement d'exponente au niveau 4, illustré à la figure 7.33. Cependant, lorsque nous avons évalué ces deux descriptions en termes de compacité, nous avons constaté que la différence de longueur de description entre ces deux analyses est négligeable – et ce malgré la prise en compte des critiques théoriques à l'égard de 3H. Nous en concluons que deux descriptions différentes structurellement peuvent être très semblables en termes d'économie de description. Dès lors que deux analyses concurrentes présentent sensiblement la même longueur de description, il devient difficile de les départager sur ce seul critère.

À l'inverse, nous constatons que la mesure de compacité est insuffisante pour évaluer à elle seule l'intégralité des analyses concurrentes pour un système flexionnel donné. La description traditionnelle 3F est en effet plus compacte en apparence que l'analyse nouvelle 3H/4H. Cependant, 3F masque entièrement les irrégularités, notamment sémantiques, que nous avons relevées au chapitre 1. Ces irrégularités doivent donc être prises en charge par des mécanismes extérieurs à la description des formes. Il n'est ainsi en rien évident que 3F sera en fin de compte plus compacte que 3H/4H, une fois ces irrégularités prises en compte par la description complète de la grammaire et du lexique. En d'autres termes, les mesures de compacité que nous avons introduites au chapitre 6 ne peuvent départager que deux descriptions qui placent la frontière entre la flexion et la dérivation au même endroit.

Afin de pouvoir réellement comparer 3F et 3H/4H, nous aurions donc besoin de pouvoir mesurer plus d'informations que celles actuellement prises en compte par notre métrique. Dans le cas de la dérivation passive, nous aurions par exemple tout intérêt à compléter notre inventaire de critères d'évaluation par une mesure de la régularité en dérivation et de son influence sur la complexité générale du système. Il serait en particulier intéressant d'explorer la question de la facilité à inférer le schème flexionnel d'un verbe par la connaissance de celui d'un autre verbe dont il est dérivé dans le cas de dérivation régulières. Cela nécessiterait cependant de pouvoir modéliser la dérivation avec au moins autant de précision que ce que nous avons tenté de faire pour la flexion avec  $\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}$ .

Mais quand bien même il existe des cas où notre mesure de compacité ne suffit pas à départager deux analyses, il n'en est pas moins que dans des cas comme celui des deux analyses concurrentes du maltais ou des quatre analyses concurrentes du français que nous avons présentées au chapitre 6, la mesure de compacité reste un moyen efficace pour choisir entre des analyses concurrentes en s'appuyant sur un principe général qui sous-tend toute démarche scientifique : celui du rasoir d'Ockham consistant à rechercher toujours l'économie descriptive.

Concernant les mesures de canonicité que nous avons appliquées aux descriptions complètes 3H/4H et 3F à la section 7.4, nous avons montré qu'elles peuvent facilement être calculées pour un système flexionnel complet à partir des propriétés particulières des flexèmes. Toutefois, il apparaît clairement que les résultats de ces mesures concernent une description particulière et non d'un système flexionnel. Les phénomènes non canoniques observés découlent en effet directement de la façon dont a été formulée l'analyse. Par exemple, nous remarquons en passant à la page 452 de ce chapitre que nous n'avons relevé, dans aucune des analyses, de cas de surdifférenciation. Or ce fait est la conséquence directe de la manière dont nous avons défini les catégories flexionnelles. Dans 3F, nous aurions pu poser une catégorie spécifique pour les verbes intransitifs, sans PTM pour le passif, le participe passé, l'adjectif verbal et le supin. Cela aurait fait des verbes déponents des verbes avec surdifférenciation pour ces formes non finies. De même, le choix d'analyser la voix comme relevant de la dérivation ou de la flexion se répercute directement dans l'existence ou non de cas de décalage morphosyntaxique dans l'analyse : si le m-passif correspond au sous-paradigme s-passif des verbes transitifs, alors leur emploi pour exprimer le s-actif relève du décalage ; si le m-passif constitue une classe flexionnelle indépendante, alors il n'y a pas lieu de parler de décalage morphosyntaxique. En revanche, cette seconde analyse induit des cas d'hétéroclise pour les verbes déponents et les passiva tantum.

Nous constatons que, comme les mesures de compacité, les mesures de canonicité ne peuvent en fin de compte être dissociées des questions de délimitation de la flexion et des autres composantes de la grammaire. Se pose notamment la question de la canonicité de la relation de dérivation entre les verbes m-actifs et m-passifs. Cette dimension n'est à ce stade pas prise en compte par nos métriques. Elles ne mesurent en particulier pour le moment pas les cas d'homophonie comme ceux que nous observons dans les analyses 3H et 4H entre les radicaux m-actifs et m-passifs dérivés. Selon les critères de canonicité des paradigmes proposés par Corbett (2007a), deux radicaux appartenant à des flexèmes différents sont canoniquement différents. Ce n'est pas le cas des radicaux RAD1 et RAD3 de AMO « *aimer* » et AMOR « *être aimé* ».

Le but de ce chapitre était d'illustrer les différentes métriques que nous avons mises en place au cours de cette thèse. Les expériences que nous y avons décrites illustrent au-delà les limites de ces métriques. En tant que métriques de la constitution et de la répartition des formes, elles ne sauraient notamment s'appliquer à des données relevant de composantes de la grammaire dépassant cette seule dimension<sup>34</sup>. Pour ce faire, il serait nécessaire de

---

34. Cette remarque n'inclut pas la morphologie dont nous avons montré dans nos expériences sur le maltais au chapitre 6 qu'elle était couverte par notre métrique.

mettre en place une modélisation complète de ces composantes voisines qui servirait alors de fondement pour le développement de nouvelles métriques plus complètes, autant pour la compacité que pour la canonicité.

Enfin, il apparaît à ce stade qu'il est difficile de tirer quelque conclusion que ce soit sur le rapport entre compacité et canonicité. Les deux métriques sont hautement dépendantes de l'analyse qui est donnée d'un système flexionnel et dont l'interaction avec les composantes voisines peut être extrêmement complexe. À défaut de pouvoir en formaliser toutes les dimensions (par exemple la dimension sémantique intervenant dans le lien entre deux formes m-active et m-passive correspondante), il n'est pour le moment pas possible d'indiquer des mesures fiables des caractéristiques plus ou moins canoniques de la langue. En revanche, les deux types de mesures nous renseignent déjà indépendamment l'une de l'autre sur différentes caractéristiques d'analyses particulières proposées.



# Conclusion générale





# Conclusion générale

Dans cette thèse nous avons cherché à proposer un modèle formel de la morphologie flexionnelle qui soit typologiquement fondé et qui permette de concevoir, d'implémenter et d'évaluer quantitativement des analyses formelles de systèmes flexionnels particuliers. Nous avons développé le modèle  $\mathcal{PARSL}$  qui s'appuie très étroitement sur les concepts de la typologie canonique (Corbett, 2003).  $\mathcal{PARSL}$  a été doté d'un système d'implémentation,  $\text{Alexina}_{\mathcal{PARSL}}$  adapté du formalisme lexical *Alexina* (Sagot, 2010). Afin de pouvoir évaluer le degré de canonicité d'une langue pour une description donnée,  $\mathcal{PARSL}$  a également été doté de mesures de canonicité. Enfin,  $\text{Alexina}_{\mathcal{PARSL}}$  a été doté d'une métrique de compacité qui permet de comparer quantitativement deux descriptions  $\mathcal{PARSL}$  implémentées en termes d'économie descriptive.

## Un modèle formel de la morphologie flexionnelle

$\mathcal{PARSL}$  est un modèle formel de la morphologie flexionnelle qui vise à formaliser directement une approche théorique de la typologie linguistique. Il s'appuie sur le paradigme de la *TYPLOGIE CANONIQUE* développé par Corbett (2003) que nous avons résumé au chapitre 3. Comme en témoigne la base de données typologiques de l'Université de Surrey<sup>35</sup>, ce paradigme typologique repose sur une longue expérience de description de données typologiquement diverses et d'étude descriptive et théorique de phénomènes linguistiques différents.

$\mathcal{PARSL}$  comporte un certain nombre d'innovations par rapports aux modèles formels de la morphologie flexionnelle existants. Comme nous l'avons indiqué à la section 4.3 du chapitre 4, l'un des apports principaux de  $\mathcal{PARSL}$  est la généralisation de la notion de

---

35. *Surrey Morphology Group Database* : base de données en ligne accessible à l'adresse <http://www.smg.surrey.ac.uk/> (valable au 11 juin 2013).

---

partitionnement des paradigmes en zones, issue de la tradition des espaces partitionnants de Pirrelli & Battista (2000) et des espaces thématiques de Bonami & Boyé (2003) et Bonami et al. (2008), à l'ensemble du mécanisme de la flexion. Ce faisant,  $\mathcal{PARSL}$  permet de rendre compte pleinement des distributions morphomiques observables dans les systèmes flexionnels, sans se limiter à des alternances radicales.

Une autre particularité de  $\mathcal{PARSL}$  est sa définition de la notion de classe flexionnelle qui n'est plus une primitive du système flexionnel mais une notion dérivée à partir des zones d'exponence (ou flexionnelles) définies. Dans  $\mathcal{PARSL}$ , les classes flexionnelles sont des associations par défaut de zones d'exponence qui maximisent par leur combinaison l'économie descriptive globale du système.

$\mathcal{PARSL}$  représente par ailleurs la flexion sous la forme d'un nombre libre de niveaux réalisationnels distincts. Il est notamment possible de représenter un niveau thématique indépendant ainsi que de faire usage d'une segmentation non seulement entre le radical et ses exposants mais également entre différents types d'exposants marqués à chaque fois par un jeu de règles de réalisation spécifique.

Cette représentation multi-niveaux permet également de généraliser la notion de classe flexionnelle à tous les niveaux sous la forme de classes réalisationnelles qui peuvent aussi bien être des classes d'exponence (flexionnelles) que des classes de radicaux ou des classes thématiques. Deux flexèmes différents peuvent ainsi partager une classe réalisationnelle à un niveau sans pour autant être associés à une même classe à un autre niveau. Cela permet de factoriser les classes définies à chaque niveau en les associant de façon croisée selon les flexèmes concernés.

$\mathcal{PARSL}$  est avant tout, du point de vue formel, un modèle qui représente de façon explicite le rapport entre la définition d'une entrée lexicale et la structure de son paradigme, et notamment la canonicité de ce dernier. Une entrée lexicale indique immédiatement dans sa structuration les caractéristiques plus ou moins canoniques de son paradigme. Ainsi, un ensemble de paquets de traits morphosyntaxiques (PTM) non standard permettra d'identifier les cas de déficience ou de surdifférenciation qui se traduiront dans la structure du paradigme par des cases manquantes ou au contraire surnuméraires. De même, des zones réalisationnelles manquantes ou surnuméraires dans le schème flexionnel de l'entrée indiqueront les cas de défektivité ou de surabondance qui se traduisent également par des cases manquantes ou surnuméraires dans le paradigme. Une liste de radicaux supplétifs ou de formes supplétives non vide indiquera la présence de cas de supplétion de radicaux, respectivement de formes. Enfin, l'association, dans un couple réalisationnel du schème flexionnel, d'une zone réalisationnelle avec une fonction de transfert différente

---

de la fonction identité marquera explicitement les cas de décalage morphosyntaxique. La structuration des entrées lexicales  $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$  a donc effectivement été conçue pour rendre compte explicitement de chacun de ces phénomènes.

Notons cependant qu'à ce stade  $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$  ne prend en compte que la partie synthétique de la morphologie flexionnelle. La réalisation des formes périphrastiques n'est pas encore couverte. Dans le cas du sous-paradigme m-passif perfectif des verbes latins, les cases comportant des réalisations périphrastiques ont ainsi été représentées de la même manière que des cases véritablement défectives. Nous ne prenons pas ici position sur la question de savoir si la périphrase doit être prise en compte par la morphologie flexionnelle ou non, ni comment. Mais nous constatons qu'une modélisation qui traite de la même manière la défectivité réelle et les cas de réalisation périphrastique n'est pas pleinement satisfaisante. Une révision du modèle pour qu'il prenne en compte la différence entre ces deux cas serait donc une extension naturelle de  $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$ .

De même,  $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$  ne traite à ce jour que la seule morphologie flexionnelle. Or, les expériences sur les descriptions concurrentes de la flexion verbale du latin au chapitre 7 ont montré que pour juger de la qualité de la représentation d'une description flexionnelle, il peut être nécessaire de pouvoir rendre compte également des propriétés dérivationnelles en particulier et constructionnelles en général du système décrit. Il n'est à ce stade pas possible de rendre compte avec  $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$  des caractéristiques constructionnelles d'entrées lexicales. Pour cela, une extension du modèle à la morphologie constructionnelle serait nécessaire.

## Un système d'implémentation de la morphologie flexionnelle

Nous avons montré sur l'exemple du maltais au chapitre 5 comment l'implémentation contribue à l'évaluation de l'adéquation d'un modèle formel un ensemble de données. L'implémentation à grande échelle garantit la cohérence et la couverture de toute analyse proposée.

$\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$  s'accompagne de son propre système d'implémentation  $\text{Alexina}_{\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}}$ . Ce système d'implémentation a été spécifiquement adapté du formalisme lexical *Alexina* (Sagot, 2010) développé pour la conception de ressources lexicales pour le traitement automatique des langues, afin qu'il représente fidèlement les primitives employées par  $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$ . Nous avons enrichi le formalisme *Alexina* par l'intégration d'un mécanisme explicite de définition des traits morphosyntaxiques permettant de calculer l'intégralité des cases des paradigmes flexionnels, et d'un mécanisme de définition d'opérations sur

---

les formes, permettant de représenter directement des opérations morphologiques non concaténatives. Enfin, nous avons également rendu plus transparente la syntaxe des grammaires Alexina afin qu'elles soient plus immédiatement utilisables par des linguistes à des fins d'implémentation de descriptions morphologiques.

Compte tenu de la nature même de la morphologie flexionnelle, garantir la couverture et la cohérence d'une analyse n'est cependant pas suffisant pour produire une analyse satisfaisante. Le problème vient plutôt du fait que, pour un système flexionnel donné, plusieurs descriptions concurrentes également cohérentes et couvrantes peuvent être réalisées. Par conséquent, de nombreux débats dans le domaine de la morphologie flexionnelle se concentrent autour de la question de l'identité de la *meilleure* des analyses proposées. Un exemple de ce type de débat est donné par les analyses concurrentes de la flexion des verbes déponents latins que nous avons présentées au chapitre 1.

En revanche, ce que signifie le qualificatif de « *meilleure analyse* » n'est en général pas clairement défini. Il était jusqu'à présent généralement admis que la meilleure analyse équivalait à la plus économique, sans pour autant que ce critère ne s'accompagne d'un moyen permettant effectivement de mesurer l'économie d'une analyse donnée par rapport à une autre. Dans cette thèse, nous avons mis en place une métrique de compacité de description permettant réellement d'évaluer quantitativement la dimension plus ou moins économique d'une description donnée par rapport à ses concurrentes. Notons que cela ne peut se faire que si des analyses concurrentes ont été formulées dans un format homogène, c'est-à-dire exprimées dans le même formalisme. À ce titre, le formalisme choisi doit donc être capable d'exprimer toutes ces analyses concurrentes, même si toutes ne feront pas nécessairement appel à l'intégralité de sa puissance d'expression.  $\mathcal{PARSL}$  a été conçu dans ce but.

## Évaluation de descriptions morphologiques concurrentes

La mesure d'évaluation de descriptions  $\mathcal{PARSL}$  concurrentes que nous avons proposée au chapitre 6 repose sur la notion de Longueur de description Minimale (Rissanen, 1984) issue de la théorie de l'information. Comme le montrent nos expériences sur le système verbal du français à la section 6.2 du chapitre 6, cette métrique permet de quantifier la différence en termes de compacité de quatre descriptions concurrentes des mêmes données et donc de les évaluer quantitativement les unes par rapport aux autres en termes d'économie descriptive.

Cette mesure de compacité permet notamment de répondre à l'une des critiques

---

majeures faites aux approches constructives de la part du paradigme des approches abstractives (Blevins, 2006, 2012) : le fait que les approches constructives rendent compte des données par des segmentations arbitraires dont il est impossible d'évaluer laquelle serait la *bonne* si tant est qu'il y en ait une. Il reste vrai que les approches constructionnelles emploient des segmentations qui sont en un premier temps nécessairement arbitraires. Par la mise en place des niveaux réalisationnels dans  $\mathbb{P}\mathbb{A}\mathbb{R}\mathbb{S}\mathbb{L}$ , nous suivons clairement cette tradition. Toutefois, notre mesure de compacité permet de contenir la part d'arbitraire dans la segmentation. Il est désormais tout à fait possible de déterminer laquelle d'entre deux segmentations produit la description la plus économique.

Par ailleurs, un avantage de cette mesure par rapport à une simple compression (Bane, 2008) de la description et des données est le fait qu'elle permet d'identifier dans la description (y compris le lexique) la part pour laquelle chacune de ses composantes contribue à la longueur globale de la description. Cela permet notamment de mesurer précisément le gain de compacité que produit une variation entre deux analyses données. Nous avons notamment vu au chapitre 7 que deux analyses du latin qui ont une compacité globale presque identique contrastent quant à la contribution de chacune des composantes de leur description à la longueur de description totale. Par ailleurs, pour deux analyses données, cette métrique permet non seulement de mesurer précisément le gain d'économie obtenu dans chaque composante de la description, mais également le poids de cette composante par rapport à la description globale.

Nos expériences au chapitre 6 ont par ailleurs contribué à illustrer à quel point l'économie descriptive est finalement une question d'équilibre optimal dans la distribution de l'information entre les différentes composantes de la grammaire. Les expériences sur le français ont en un premier temps montré que la morphologie flexionnelle n'est pas un îlot isolé mais que la conception d'une description idéalement compacte nécessite une distribution optimale de l'information, et en particulier des irrégularités, entre le lexique et la grammaire. Nos expériences sur le maltais ont par ailleurs montré que la composante morphologique de la grammaire peut également prendre en charge certaines dimensions de la description. L'emploi de règles morphologiques plus ou moins nombreuses et plus ou moins complexes peut ainsi alléger ou non la taille de la description flexionnelle. Dans le cas du maltais, la compacité semble être plus grande dès lors que l'on limite la part d'information contenue dans la composante flexionnelle à la seule morphologie vraiment autonome. Compte tenu de la nature de la morphologie, dont les règles sont à chaque fois définies pour une (sous-)classe d'entrées lexicales, par rapport à la morphologie dont les règles s'appliquent de façon systématique, ce résultat est par ailleurs attendu : la

---

systématicité des règles morphologiques permet une factorisation plus importante de l'information et mène donc à une plus grande économie descriptive globale. Les résultats de notre métrique qui, à ce stade, prend en compte la part d'information contenue dans le lexique, la morphologie et la grammaire flexionnelle, semblent ainsi justifier la conception modulaire articulée de la grammaire que nous avons proposée à la section 2.2 du chapitre 2.

Dans son état actuel, la métrique proposée a cependant un certain nombre de limites. Dans le chapitre 7, nous avons comparé une analyse de la flexion verbale latine qui traite la passivation comme un phénomène dérivationnel à une analyse qui la considère comme un phénomène flexionnel. Il s'avère que l'analyse flexionnelle est plus compacte selon cette métrique. Or, les arguments qui nous avaient fait préférer au chapitre 1 une analyse dérivationnelle de la passivation en latin relevaient de l'imprédictibilité sémantique qu'elle comporte. Cette imprédictibilité, comme la complexité qu'elle induit, est cependant complètement masquée dans les analyses implémentées sur lesquelles reposent la mesure de compacité. Nous touchons là à une limite de notre métrique. Dans la mesure où `PARSL` ne modélise pas encore la morphologie constructionnelle, notre mesure ne saurait prendre en compte la part d'information comprise dans cette composante. Il sera pour cela nécessaire de pouvoir formaliser et implémenter les entrées lexicales au niveau lexématique, et non pas seulement flexématique, et les liens constructionnels existant entre elles. Cette formalisation devra notamment prendre en compte les régularités en termes de sémantique, de valence, de changements de structure argumentale et de propriétés constructionnelles avant de pouvoir servir de fondement à une métrique de compacité sachant départager deux analyses différant dans leur définition des composantes morphologiques.

Par ailleurs, il serait naturellement intéressant de pouvoir comparer les résultats de la métrique que nous avons proposée pour l'évaluation de descriptions morphologiques avec d'autres métriques, qu'il s'agisse de métriques permettant d'évaluer le fonctionnement de la morphologie pour un locuteur ou de métriques purement computationnelles sur le temps d'exécution nécessaire pour la production des formes par une machine. Nous ne nous sommes toutefois pas penché sur ces questions pour le moment.

## Un modèle formel pour la typologie

`PARSL` est aujourd'hui un modèle formel qui permet, grâce à un système d'implémentation et à une métrique de compacité dédiés, de modéliser et évaluer une description flexion-

---

nelle particulière par rapport à d'autres descriptions flexionnelles concurrentes. Mais le but de  $\mathcal{PARSL}$  reste avant tout d'être un modèle formel à l'intention de la typologie linguistique.

L'objectif de la typologie linguistique réside dans la comparaison de systèmes linguistiques divers. Avec  $\mathcal{PARSL}$ , nous avons cherché à mettre en place un outil au service de la comparaison typologique. La façon dont il a été conçu, en s'appuyant directement sur un paradigme typologique existant, reprend des concepts qui ont déjà fait leur preuve dans la description de langues diverses. La formalisation de ces concepts vise à rendre plus facile la comparaison entre des systèmes linguistiques différents en favorisant leur utilisation rigoureuse par le biais de la modélisation. L'implémentation de cette modélisation permet de surcroît de vérifier sa cohérence, et donc celle de l'analyse. Le fait que le formalisme d'implémentation  $\text{Alexina}_{\mathcal{PARSL}}$  permette par ailleurs de prendre en compte facilement un grand nombre de données offre enfin l'assurance que toutes les données que l'on cherchait à comparer sont effectivement bien couvertes par l'analyse.

Enfin, les mesures de canonicité développées dans  $\mathcal{PARSL}$  ont permis de doter l'approche canonique de la morphologie flexionnelle de mesures quantitatives pour la description des langues à partir de descriptions implémentables et évaluables.  $\mathcal{PARSL}$  est ainsi un outil de description qui permet de représenter des données dans un format homogène afin de faciliter la comparaison de différents systèmes linguistiques. La qualité de la description proposée pour chacun d'entre eux peut par ailleurs être évaluée indépendamment en termes d'économie descriptive. Il va cependant de soi que la comparaison ne fait sens que sur des descriptions de qualité comparable et couvrant une proportion comparable de chacune des langues comparées, notamment quant à la taille du lexique et au nombre de catégories flexionnelles couvertes. La comparaison sera d'autant plus réaliste que la description sera complète.

Nous avons toutefois constaté par les expériences sur le latin au chapitre 7 que l'identification des propriétés canoniques d'une langue ne va pas de soi. Les phénomènes que l'on observe dans les langues dépendent directement de la manière dont les données ont été analysées, et le plus souvent plusieurs analyses restent possibles qui mènent à des mesures de canonicité différentes. Cette ambiguïté entre la définition des phénomènes mériterait probablement d'être étudiée en tant que telle dans l'avenir.



---

## Vers une morphologie de terrain outillée

Comme nous l'avons indiqué ci-dessus, la comparaison typologique repose sur la disponibilité de données extensives dont la collecte peut nécessiter un retour sur le terrain. Dans la mesure où  $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$  a été doté d'un formalisme d'implémentation, Alexina $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$ , il peut servir d'outil pour la linguistique de terrain. Alexina $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$  repose sur le formalisme lexical Alexina développé pour le traitement automatique des langues (TAL) et pourvu d'outils d'enrichissement automatique de lexiques (Sagot, 2005). Ces outils permettent à partir d'une description nouvelle d'extraire rapidement de nouvelles entrées lexicales à partir de corpus bruts<sup>36</sup> et de construire un lexique de formes fléchies à partir de données nouvelles. Les formes générées par le fléchisseur d'Alexina $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$  devraient pouvoir être utiles pour des études de terrain où il serait ainsi plus facile d'avoir accès aux formes que l'on chercherait à faire vérifier par des locuteurs. Le fait qu'Alexina $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$  ait été adapté pour être à la fois plus facile d'utilisation et plus cohérent avec les réalités linguistiques observables devrait rendre plus accessible au linguiste de terrain les outils dont est pourvu le formalisme Alexina. Par ailleurs, Alexina est dotée d'une interface de validation permettant facilement l'évaluation des formes produites par l'analyse (Sagot, 2005 ; Sagot et al., 2011), simplifiant encore le travail de vérification sur le terrain.

De plus, Alexina est avant tout un formalisme utilisé en TAL. Or, le traitement de grandes quantités de données comme celui qui peut être effectué en TAL serait d'une grande utilité pour l'étude linguistique. Une extension des travaux décrits dans cette thèse serait ainsi de renforcer encore les liens entre les différents domaines de la linguistique que sont la typologie linguistique, la linguistique théorique et formelle et la linguistique informatique, afin de concevoir des analyses étendues pour des langues peu décrites mais pour lesquelles des corpus bruts sont disponibles, et ce notamment dans l'objectif de travaux de comparaison typologique à grande échelle. En retour, un tel travail donnerait à la communauté du TAL accès à des données rares susceptibles de nourrir des systèmes d'analyse semi-automatique.

---

36. Ces outils n'ont cependant pas encore été adaptés au nouveau formalisme Alexina $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$ .

# Annexes



# Implémentations des descriptions morphologiques

## A.1 Le système verbal du latin

### A.1.1 Formalisation des règles de réalisation en *Paradigm Function Morphology*

Afin de rendre compte de notre analyse du système verbal latin, nous indiquons ici, au format *Paradigm Function Morphology*, les règles de réalisation utilisées pour la génération des formes verbales synthétiques. Une représentation en Alexina (Sagot, 2010) est donnée à la section suivante.

Les données présentées dans la section 1.1.1 et dans les tableaux 1.17, 1.18, 1.21 et 1.22, permettent d'établir les classes de positions indiquées dans les tableaux 1.25 à 1.27 au chapitre 1 pour les classes flexionnelles A, B, C et D respectivement.

En guise de modélisation des verbes déponents et du passif en latin, nous pouvons ainsi écrire les règles de réalisations en format *Paradigm Function Morphology* (PFM) ci-dessous. Ces règles font intervenir un inventaire de traits morphosyntaxiques que nous avons défini dans le tableau 1.24. Les différents types de verbes forment leurs paradigmes en utilisant les zones définies plus haut. Pour l'application des règles ci-dessous, un lexème donné sélectionnera donc la règle qui correspond à la classe flexionnelle à laquelle appartient la zone selon laquelle il se fléchit compte tenu des traits à exprimer.

Dans la mesure où PFM permet également de définir des super-classes pour les classes de lexèmes, nous définissons les quatre super-classes indiquées dans la figure A.1.

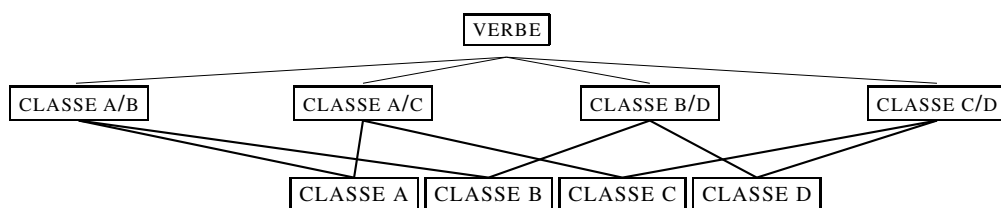


FIGURE A.1 : Représentation des super-classes verbales correspondant à l'analyse du chapitre 1

Ainsi, un verbe intransitif standard de classe A sélectionnera exclusivement des règles pour des lexèmes de type  $X_{V-A}$  ou appartenant à une des super-classes contenant A, cf. V-A/C. Un verbe à valeur *s*-passive standard de classe B sélectionnera exclusivement des règles pour lexèmes de type  $X_{V-B}$  ou appartenant à une des super-classes contenant B, en cf. V-B/D. Un verbe déponent standard de classe B+A sélectionnera toutes les règles pour  $X_{V-B}$  plus les règles pour  $X_{V-A}$  pour les règles réalisant les formes de la zone I-ZONE3 de la classe A (définie plus haut), ou alors des règles appartenant une des super-classes compatibles.

Les règles PFM s'organisent selon des blocs de règles qui s'appliquent dans l'ordre, ici en l'occurrence le bloc 1 avant le bloc 2. Les règles dites *porte-manteau* sont prioritaires sur les autres. Elles s'appliquent donc nécessairement avant les blocs simples (1 et 2).

Notons par ailleurs que nous supposons dans cette analyse l'existence des règles morphologiques aux frontières de morphes.

- a-o → o
- bi-o → bo
- be-or → bor
- bi-nt → bunt
- a-a → e

#### Porte-manteau 1-2

$X_{V-A/C}$ , {temps *prs*, aspect *ipfv*, mode *inf*} → Xre

$X_{V-B/D}$ , {temps *prs*, aspect *ipfv*, mode *inf*} → Xri

$X_{V-A/C}$ , {personne 2, nombre *sg*, temps *prs*, aspect *ipfv*, mode *imp*} → X  
 $X_{V-B/D}$ , {personne 2, nombre *sg*, temps *prs*, aspect *ipfv*, mode *imp*} → Xre  
 $X_{V-B/D}$ , {personne 2, nombre *pl*, temps *prs*, aspect *ipfv*, mode *imp*} → Xmini  
 $X_{V-A/C}$ , {personne 3, nombre *pl*, temps *fut*, aspect *ipfv*, mode *imp*} → Xnto  
 $X_{V-B/D}$ , {personne 2, nombre *sg*, temps *fut*, aspect *ipfv*, mode *imp*} → Xtor  
 $X_{V-B/D}$ , {personne 3, nombre *sg*, temps *fut*, aspect *ipfv*, mode *imp*} → Xtor  
 $X_{V-B/D}$ , {personne 3, nombre *pl*, temps *fut*, aspect *ipfv*, mode *imp*} → Xntor

$X_{V-A/C}$ , {temps *pst*, aspect *pfv*, mode *inf*} → Xisse  
 $X_{V-A/C}$ , {personne 1, nombre *sg*, temps *pst*, aspect *pfv*, mode *ind*} → Xi  
 $X_{V-A/C}$ , {personne 2, nombre *sg*, temps *pst*, aspect *pfv*, mode *ind*} → Xisti  
 $X_{V-A/C}$ , {personne 2, nombre *pl*, temps *pst*, aspect *pfv*, mode *ind*} → Xistis

$X_V$ , {nom.nombre *sg*, temps *prs*, mode *ptcp*} → Xns

#### Bloc 1

$X_V$ , {temps *prs*, aspect *ipfv*, mode *sbjv*} → Xa  
 $X_{V-A/B}$ , {nom.temps *pst*, aspect *ipfv*, mode *ind*} → Xba  
 $X_{V-C/D}$ , {nom.temps *pst*, aspect *ipfv*, mode *ind*} → Xeba  
 $X_V$ , {temps *pst*, aspect *ipfv*, mode *sbjv*} → Xre  
 $X_{V-A/B}$ , {temps *fut*, aspect *ipfv*, mode *ind*} → Xbi  
 $X_{V-C/D}$ , {temps *fut*, aspect *ipfv*, mode *ind*} → Xe  
 $X_V$ , {temps *fut*, aspect *ipfv*, mode *imp*} → Xto

$X_{V-A/C}$ , {temps *pst*, aspect *pfv*, mode *ind*} → Xi  
 $X_{V-A/C}$ , {temps *pst*, aspect *pfv*, mode *ind*} → Xera  
 $X_{V-A/C}$ , {temps *pst*, aspect *pfv*, mode *sbjv*} → Xisse  
 $X_{V-A/C}$ , {temps *fut*, aspect *pfv*, mode *ind*} → Xeri  
 $X_{V-A/C}$ , {temps *prs*, aspect *pfv*, mode *sbjv*} → Xeri

$X_V$ , {temps *prs*, mode *ptcp*} → Xnt  
 $X_V$ , {temps *fut*, mode *ptcp*} → Xur  
 $X_V$ , {mode *sup*} → X  
 $X_V$ , {mode *grd*} → Xnd

$X_V, \{\text{mode } grv\} \rightarrow X_{nd}$

## Bloc 2

$X_{V-A/C}, \{\text{personne } 1, \text{ nombre } sg, \text{ mode } ind\} \rightarrow X_o$

$X_{V-B/D}, \{\text{personne } 1, \text{ nombre } sg, \text{ aspect } ipfv, \text{ mode } ind\} \rightarrow X_{or}$

$X_{V-A/C}, \{\text{personne } 1, \text{ nombre } sg, \text{ mode } sbjv\} \rightarrow X_m$

$X_{V-A/C}, \{\text{personne } 1, \text{ nombre } sg, \text{ temps } pst, \text{ mode } ind\} \rightarrow X_m$

$X_{V-B/D}, \{\text{personne } 1, \text{ nombre } sg, \text{ aspect } ipfv\} \rightarrow X_r$

$X_{V-A/C}, \{\text{personne } 2, \text{ nombre } sg\} \rightarrow X_s$

$X_{V-B/D}, \{\text{personne } 2, \text{ nombre } sg, \text{ aspect } ipfv\} \rightarrow X_{ris}$

$X_{V-A/C}, \{\text{personne } 3, \text{ nombre } sg\} \rightarrow X_t$

$X_{V-B/D}, \{\text{personne } 3, \text{ nombre } sg, \text{ aspect } ipfv\} \rightarrow X_{tur}$

$X_{V-A/C}, \{\text{personne } 1, \text{ nombre } pl\} \rightarrow X_{mus}$

$X_{V-B/D}, \{\text{personne } 1, \text{ nombre } pl, \text{ aspect } ipfv\} \rightarrow X_{mur}$

$X_{V-A/C}, \{\text{personne } 2, \text{ nombre } pl\} \rightarrow X_{tis}$

$X_{V-B/D}, \{\text{personne } 2, \text{ nombre } pl, \text{ aspect } ipfv\} \rightarrow X_{mini}$

$X_{V-A/C}, \{\text{personne } 3, \text{ nombre } pl\} \rightarrow X_{nt}$

$X_{V-B/D}, \{\text{personne } 3, \text{ nombre } pl, \text{ aspect } ipfv\} \rightarrow X_{ntur}$

$X_{V-A/C}, \{\text{personne } 2, \text{ nombre } pl, \text{ aspect } ipfv, \text{ mode } imp\} \rightarrow X_{te}$

Bloc 2 — règles supplémentaires pour la déclinaison des participes Les participes empruntent leurs désinences casuelles aux classes flexionnelles adjectivales. Nous ne détaillons pas ici le processus de réalisation de l'intégralité des formes participiales. Le lecteur peut néanmoins se référer à la section A.1.2 pour lire la description complète au format Alexina.

### A.1.2 Implémentation native en Alexina

Cette implémentation en Alexina<sub>PARSLI</sub><sup>1</sup> peut être trouvée en ligne sur la page personnelle de l'auteur. Elle comporte les règles de réalisation données dans la section

---

1. Voir le chapitre 5 pour une description de l'implémentation du modèle PARSLI développée dans ce formalisme adapté du formalisme lexical Alexina (Sagot, 2010).

précédente et correspondant à l'analyse des verbes latins présentée au chapitre 1, ainsi que les règles de formation de radicaux correspondantes. Elle correspond à l'implémentation de l'analyse 3H décrite à la section 7.1 du chapitre 7.

### A.1.3 Implémentation de la variante de l'analyse en trois niveaux avec hétéroclise radicale en Alexina

L'implémentation présentée dans cette section est une variante de celle de la section précédente. La différence entre les deux analyses réside dans la stipulation de cas d'hétéroclise au niveau radical dans l'analyse ci-dessous. Cette différence n'affecte que la définition des classes réalisationnelles au niveau radical et des schèmes flexionnels. La description complète est accessible sur la page personnelle de l'auteur sous le nom de fichier 3H-var2.

### A.1.4 Implémentation de l'analyse *Network Morphology* en DATR

L'implémentation complète en DATR de l'analyse des déponents latins proposée par Hippiusley (2007) est disponible en ligne sur le site du livre *Network Morphology* (Brown & Hippiusley, 2012) à l'adresse [http://networkmorphology.as.uky.edu/sites/default/files/ch5\\_latin.dtr](http://networkmorphology.as.uky.edu/sites/default/files/ch5_latin.dtr)<sup>2</sup>.

### A.1.5 Implémentation de l'analyse à quatre niveaux du système verbal latin en Alexina<sup>PARRSLI</sup>

L'implémentation dans cette section correspond à l'analyse 4H de la flexion verbale latine avec deux niveaux d'exponence distincts présentée à la section 7.2 du chapitre 7. Le lexique fléchi est le même que celui de l'analyse à un seul niveau d'exponence dont l'implémentation est décrite à la section A.1.2. Les deux analyses sont également identiques aux niveaux radical et thématique ainsi que dans leur inventaire de traits et, pour la plus grande partie, dans la spécification des règles morphologiques. L'analyse est disponible sur la page personnelle de l'auteur sous le nom de 4h-orig.

---

2. (valable au 11 juin 2013).



### A.1.6 Implémentation de la variante de l'analyse en quatre niveaux avec hétéroclise radicale en Alexina

L'implémentation présentée dans cette section est une variante de celle de la section précédente. La différence entre les deux analyses réside dans la stipulation de cas d'hétéroclise au niveau radical dans l'analyse ci-dessous. Cette différence n'affecte que la définition du niveau radical et des schèmes flexionnels. La description complète est accessible sur la page personnelle de l'auteur sous le nom de 4H-var.

### A.1.7 Implémentation de l'analyse traditionnelle en trois niveaux en Alexina

L'implémentation de l'analyse traditionnelle 3F décrite au chapitre 7 est disponible sur le site de l'auteur sous le nom 3F. Elle se distingue des autres analyses des sections précédentes par le fait qu'elle considère le passif comme un attribut exprimé par la flexion. Dans cette analyse, les verbes intransitifs sont déficients pour le sous-paradigme passif. Les verbes déponents sont également déficients pour le sous-paradigme passif, hormis pour les formes non finies. Ils présentent par ailleurs un décalage morphosyntaxique en ce sens qu'ils emploient des zones flexionnelles du passif pour réaliser leurs formes finies actives. Cette implémentation correspond à une implémentation d'une description proche de l'analyse implémentée en DATR de Hippisley (2007).

## A.2 Le premier binyan du système verbal à base sémitique du maltais

### A.2.1 Implémentation de l'analyse morphologique de la flexion du premier binyan en Alexina<sup>PARSLI</sup>

La grammaire implémentée de notre analyse morphologique du premier binyan des verbes à base sémitique du maltais que nous avons donnée au chapitre 5 est disponible sur le site de l'auteur sous le nom de maltese-m.

## A.2.2 Implémentation de l'analyse morphologique de la flexion du premier binyan en Alexina<sup>PARSLI</sup>

La grammaire implémentée de notre analyse morphologique du premier binyan des verbes à base sémitique du maltais que nous avons donnée au chapitre 6 est disponible sur la page personnelle de l'auteur sous le nom `Maltese-p`. Elle se différencie de celle de la précédente par son emploi de règles morphologiques (`sandhi`) plus nombreuses et plus complexes, mais un nombre bien plus réduit de schèmes flexionnels et de classes réalisationnelles.



## Schémes flexionnels distincts dans *Leffla*

Nous indiquons ci-dessous le nombre de schémas flexionnels distincts employés dans nos trois analyses de la flexion verbale du latin telles que nous les avons décrites au chapitre 7. Nous indiquons également pour chaque schème le nombre de flexèmes qui l'emploient.

Les schémas ont été nommés pour donner les informations suivantes :

- *a, e, i, c* ou *ic* indique pour chaque schème s'il comporte un thème en *a, e, i* ou *c* ou un thème mixte en *i* et *c* ;
- les A, B, C et D en lettres capitales correspondent aux classes d'exponence : A pour  $Z_{v-A}^{exp}$ , B pour  $Z_{v-B}^{exp}$ , C pour  $Z_{v-C}^{exp}$  et D pour  $Z_{v-D}^{exp}$ . Les combinaisons de type AB ou CD sont des cas d'hétéroclise entre les classes correspondantes ;
- le reste du schème indique à chaque fois la classe de radicaux employée : par exemple pour REG-V-T la classe  $Z_{REG-V-T}^s$ .

### B.1 Analyses 3H et 4H

358	v-aB-REG-V-T	62	v-eB-REG-V-T
352	v-aA-REG-V-T	54	v-icC-REG-EQ-T
253	v-cD-REG-V-T	52	v-cC-REG-EQ-S
92	v-cC-REG-EQ-T	48	v-cC-SUPPL
85	v-cC-REG-X-T	41	v-iD-REG-V-T
83	v-eA-REG-U-T	37	v-cC-REG-V-T
76	v-cD-REG-EQ-S	36	v-cC-H-S-T
62	v-aAB-REG-V-T	35	v-cD-REG-S-SS

31	v-icD-REG-EQ-T	5	v-eA-REG-EQ-T
28	v-cC-H-S-S	5	v-iD-REG-EQ-S
27	v-cCD-REG-V-T	4	v-eA-H-EQ-SS
25	v-icD-REG-V-T	4	v-aBA-REG-U-T
25	v-iC-REG-V-T	4	v-iC-REG-EQ-T
20	v-eA-REG-V-T	4	v-iC-REG-X-T
19	v-cC-REG-U-T	4	v-eB-REG-S-SS
18	v-cC-REG-S-SS	4	v-icC-H-SS-SS
17	v-cD-REG-EQ-T	4	v-cCD-REG-EQ-S
17	v-aBA-REG-V-T	3	v-cC-H-X-EQ
16	v-feror-SUPPL	3	v-eA-SUPPL
16	v-fero-SUPPL	3	v-cC-H-EQ-EQ
14	v-cD-H-S-T	3	v-eAB-REG-U-T
13	v-sum-SUPPL	3	v-cDC-REG-EQ-S
13	v-cD-REG-UL	2	v-aA-REG-U-T
12	v-eo-SUPPL	2	v-aA-REG-EQ-T
12	v-iCD-REG-V-T	2	v-eB-REG-U-T
11	v-eB-REG-EQ-S	2	v-icC-SUPPL
11	v-cCD-REG-S-SS	2	v-cCD-REG-X-T
10	v-eA-H-EQ-VT	2	v-cCD-H-S-S
10	v-eB-H-EQ-VT	2	v-iC-SUPPL
9	v-cC-H-SS-SS	2	v-eAB-H-SS-SS
9	v-eA-H-S-S	1	v-iC-H-S-T
7	v-cC-REG-UL	1	v-icC-REG-EQ-S
7	v-eA-REG-EQ-S	1	v-iCD-REG-U-T
7	v-eor-SUPPL	1	v-iDC-REG-EQ-T
6	v-cC-H-SS-T	1	v-iCD-REG-EQ-S
6	v-eAB-REG-V-T	1	v-cD-REG-U-T
6	v-cD-REG-X-X	1	v-cC-H-EQ-SS
6	v-iC-REG-U-T	1	v-icC-REG-V-T
6	v-cCD-REG-X-X	1	v-eAB-H-S-S
6	v-cC-REG-X-X	1	v-eA-REG-X-T
6	v-aA-SUPPL	1	v-eAB-REG-S-SS
6	v-cD-H-EQ-EQ	1	v-icC-REG-U-T
6	v-icD-REG-S-SS	1	v-icC-H-EQ-SS
5	v-iC-H-S-S		

## B.2 Analyse 3F

352	v-aA-REG-V-T	5	v-iC-H-S-S
93	v-cC-REG-EQ-T	5	v-eA-REG-EQ-T
86	v-cC-REG-X-T	4	v-eA-H-EQ-SS
83	v-eA-REG-U-T	4	v-aBA-REG-U-T
62	v-aAB-REG-V-T	4	v-iC-REG-EQ-T
54	v-icC-REG-EQ-T	4	v-iC-REG-X-T
50	v-cC-REG-EQ-S	4	v-icC-H-SS-SS
48	v-cC-SUPPL	4	v-cCD-REG-EQ-S
38	v-cC-REG-V-T	3	v-cC-H-X-EQ
36	v-cC-H-S-T	3	v-eA-SUPPL
28	v-cC-H-S-S	3	v-icC-REG-V-T
27	v-cCD-REG-V-T	3	v-cC-H-EQ-EQ
25	v-iC-REG-V-T	3	v-eAB-REG-U-T
20	v-eA-REG-V-T	3	v-cDC-REG-EQ-S
19	v-cC-REG-U-T	2	v-aA-REG-U-T
18	v-cC-REG-S-SS	2	v-aA-REG-EQ-T
17	v-aBA-REG-V-T	2	v-cC-H-EQ-SS
16	v-fero-SUPPL	2	v-icC-SUPPL
13	v-sum-SUPPL	2	v-cCD-REG-X-T
12	v-eo-SUPPL	2	v-cCD-H-S-S
12	v-iCD-REG-V-T	2	v-iC-SUPPL
11	v-cCD-REG-S-SS	2	v-eAB-H-SS-SS
10	v-eA-H-EQ-VT	1	v-iC-H-S-T
9	v-cC-H-SS-SS	1	v-icC-REG-EQ-S
9	v-eA-H-S-S	1	v-iCD-REG-U-T
7	v-cC-REG-UL	1	v-iDC-REG-EQ-T
7	v-eA-REG-EQ-S	1	v-iCD-REG-EQ-S
6	v-cC-H-SS-T	1	v-eAB-H-S-S
6	v-eAB-REG-V-T	1	v-eA-REG-X-T
6	v-iC-REG-U-T	1	v-eAB-REG-S-SS
6	v-cCD-REG-X-X	1	v-icC-REG-U-T
6	v-cC-REG-X-X	1	v-icC-H-EQ-SS
6	v-aA-SUPPL		



# Interaction avec une théorie indépendante de la notion de traits morphosyntaxiques

## C.1 Une théorie canonique des hiérarchies de traits

En plus des caractéristiques formelles indiquées au chapitre 3, Corbett (2010) indique un critère distinctif supplémentaire des systèmes flexionnels canoniques. Canoniquement, les traits réalisés par les formes d'un paradigme correspondent à une hiérarchie régulière. Par exemple, les noms du latin marquent deux nombres (SG et PL) et pour chacun de ces nombre une distinction entre quatre cas (NOM, ACC, GEN, DAT, ABL).

Parfois, les langues affichent cependant des hiérarchies de traits qui ne se structurent pas de façon aussi régulière. Les cas que nous avons vus jusqu'ici concernent la neutralisation, ainsi que la surdifférenciation de paquets de traits morphosyntaxiques. Ci-dessous, nous décrivons quelques cas d'irrégularités observables par l'intermédiaire des relations entre traits.

### C.1.1 Sous-spécification et neutralisation

On parle de SOUS-SPÉCIFICATION dès lors que pour un (ensemble de) lexèmes un trait morphosyntaxique ne distingue pas entre (au moins) deux valeurs.

Un cas de sous-spécification peut être illustré par le comportement de certains adjectifs en français. À l'écrit, les adjectifs français distinguent entre deux valeurs pour le nombre et deux valeurs pour le genre, cf. le paradigme de BRUN dans le tableau C.1.

Or, certains adjectifs comme les adjectifs en *-ique* peuvent être conçus comme étant sous-spécifiés en genre (cf. tableau C.2), tandis que d'autres, comme l'adjectif GRIS, seront sous-



	MASC	FEM
SG	<i>brun</i>	<i>brune</i>
PL	<i>bruns</i>	<i>brunes</i>

TABLE C.1 : Le paradigme de l'adjectif BRUN

spécifiés en nombre au masculin (cf. tableau C.3).

	MASC	FEM
SG	<i>pratique</i>	<i>pratique</i>
PL	<i>pratiques</i>	<i>pratiques</i>

TABLE C.2 : Le paradigme de l'adjectif PRATIQUE

	MASC	FEM
SG	<i>gris</i>	<i>grise</i>
PL	<i>gris</i>	<i>grises</i>

TABLE C.3 : Le paradigme de l'adjectif GRIS

En  $\mathcal{PARSL}$ , ce type de phénomène sera représenté par les mêmes mécanismes que pour les cas de syncrétismes.

Parfois, au-delà de la simple sous-spécification de valeurs pour un trait donné, c'est un trait complet qui peut se retrouver non distingué dans une partie d'un paradigme donné. Le phénomène non canonique correspondant est ce qu'il convient d'appeler la NEUTRALISATION (de traits).

Un cas de neutralisation est celui du genre dans le sous-paradigme pluriel des verbes maltais. Les verbes maltais marquent en effet le genre (masculin ou féminin) à la troisième personne du singulier, mais cette distinction de genre est neutralisée au pluriel, autant au perfectif qu'à l'imperfectif (cf. tableau C.4).

Ce phénomène est généralisé à travers l'intégralité des paradigmes verbaux du maltais. Il n'appelle ainsi pas de modélisation particulière en  $\mathcal{PARSL}$ .

### C.1.2 Surdifférenciation de valeurs et de traits

L'inverse de la sous-spécification est la SURDIFFÉRENCIATION DE VALEURS DE TRAITS. Comme nous l'avons indiqué plus haut, on parle de surdifférenciation quand un lexème

	PFV	IPFV
1.SG	<i>ktibt</i>	<i>nikteb</i>
2.SG	<i>ktibt</i>	<i>tikteb</i>
3.SG.M	<i>kiteb</i>	<i>jikteb</i>
3.SG.F	<i>kitbet</i>	<i>tikteb</i>
1.PL	<i>ktibna</i>	<i>niktebu</i>
2.PL	<i>ktibtu</i>	<i>tiktebu</i>
3.PL	<i>kitbu</i>	<i>jiktebu</i>

TABLE C.4 : Le paradigme verbal KITEB « lire » (maltais)

donné distingue plus de valeurs pour un trait donné que ce qui serait attendu compte tenu des valeurs distinguées par les flexèmes de la classe morphosyntaxique à laquelle il appartient. Cette surdifférenciation introduit ainsi également une irrégularité dans les hiérarchies de traits de la langue en question. Elle se manifestera dans une entrée  $\mathbb{P}\mathbb{R}\mathbb{S}\mathbb{L}$  par un paquet de traits morphosyntaxiques surnuméraire dans l'ensemble des paquets de traits morphosyntaxiques d'un flexème particulier par rapport à l'ensemble des paquets de traits morphosyntaxiques réalisés par les flexèmes appartenant canoniquement à la catégorie flexionnelle du flexème en question. Nous avons montré un exemple de surdifférenciation de valeurs de traits en ce sens pour les noms slovaques distinguant au singulier une valeur du trait CAS supplémentaire, à savoir le cas *vocatif*.

Tout comme il y avait la neutralisation de traits qui correspondait à un cas plus extrême de sous-spécification, il arrive qu'au delà de la surdifférenciation de valeurs nous observions une surdifférenciation de traits, c'est-à-dire un cas où un sous-paradigme distingue plus de traits que le reste du paradigme. On trouve un exemple de cela en polonais, où (en plus du temps, de l'aspect, du mode et de la voix) les verbes distinguent le nombre et la personne. Dans le sous-paradigme passé, les verbes du polonais affichent en plus une distinction de genre au singulier et une opposition d'ANIMACITÉ au pluriel, en l'occurrence entre le *masculin animé* et les « autres », cf. tableau C.5.

### C.1.3 Difficultés liées à l'interprétation de la non-canonieité dans les hiérarchies de traits

Notons qu'il n'est pas toujours facile de déterminer lequel des phénomènes non canoniques est effectivement affiché par des données particulières. Nous avons ainsi cité les verbes maltais (tableau C.4) comme un exemple de neutralisation de traits, le trait de

		SG	PL
PRS	1	<i>nios-ę</i>	<i>niesie-my</i>
	2	<i>niesie-sz</i>	<i>niesie-cie</i>
	3	<i>niesie</i>	<i>nios-ą</i>
PST	1.M.ANIM	<i>nios-t-em</i>	<i>nieś-t-i-śmy</i>
	1.M.INANIM		<i>nios-t-y-śmy</i>
	1.N		
	1.F	<i>nios-t-a-m</i>	
	2.M.ANIM	<i>nios-t-eś</i>	<i>nieś-t-i-ście</i>
	2.M.INANIM		<i>nios-t-y-ście</i>
	2.N		
	2.F	<i>nios-t-a-ś</i>	
	3.M.ANIM	<i>niós-t</i>	
	3.M.INANIM		<i>nios-t-y</i>
	3.N	<i>nios-to</i>	
	3.F	<i>nios-t-a</i>	

TABLE C.5 : Le verbe polonais NIEŚĆ « porter » au présent et au prétérit (Grappin, 1985)

genre différencié à la troisième personne au singulier étant neutralisé au pluriel. Une autre interprétation de ces mêmes données serait de dire qu'il y a surdifférenciation en genre à la troisième personne du singulier ou encore que les troisièmes personnes du pluriel sont syncrétiques en genre. Trancher ce type de question n'est pas possible à l'aide d'un paradigme isolé. Pour savoir quel est le comportement habituel d'une langue donnée (par exemple le fait de comprendre une distinction de genre à la troisième personne ou non) et en déduire le type de phénomène non canonique affiché par un paradigme particulier, il est nécessaire de le replacer dans son système morphologique intégral. Ce fait rejoint ce que nous défendons au chapitre 1, à savoir la nécessité d'une modélisation intégrale d'un système morphologique donné et le besoin de prendre en compte une grande quantité de données pour pouvoir présenter une analyse cohérente d'un phénomène dans le système en question.

La difficulté à dissocier les phénomènes semble plus grande encore si l'on compare les cas de surdifférenciation et de neutralisation au syncrétisme. En effet, l'on ne voit pas bien comment il serait possible de décider formellement si les adjectifs du français comme PRATIQUE présentent des syncrétismes systématiques entre les formes du masculin et du féminin ou s'il s'agit d'un cas de neutralisation du genre. Cette distinction semble en fin de compte davantage de relever deux points de vue différents sur les données, l'un à partir des inventaires de traits morphosyntaxiques formellement distingués, l'autre à partir de la

structure formelle brute du paradigme.

Ainsi, si l'implémentation de hiérarchies de traits est immédiate en *Network Morphology* par la mise en correspondance de ladite hiérarchie avec les hiérarchies morphologiques et lexicales, ni PFM, ni  $\mathcal{PARSL}$  ne comportent en tant que tel de système indépendant indiquant les hiérarchies de traits. En revanche, il serait envisageable de mettre en correspondance les propriétés paradigmatiques identifiées et une hiérarchie de traits externe. En ce sens, il serait alors question pour les hiérarchies de traits d'une caractérisation du système flexionnel ne relevant pas de la structure paradigmatique, objet d'une modélisation  $\mathcal{PARSL}$ , mais d'une dimension orthogonale interfaçable avec la première.

## C.2 Représentation formelle de la neutralisation et de la surdifférenciation de traits

Une de spécificités de  $\mathcal{PARSL}$  est d'encoder explicitement l'articulation entre la définition d'une entrée lexicale et la forme du paradigme. Cette caractéristique se traduit notamment par le fait qu'une entrée lexicale est définie par un certain nombre de caractéristiques dont les schèmes sont indiqués plus haut.

Une entrée lexicale comporte également parmi ces éléments définitoires la catégorie morphologique  $\mathcal{C}$  à laquelle elle appartient ainsi que l'ensemble des PAQUETS DE TRAITS MORPHOSYNTAXIQUES  $\Sigma_i$  que ses formes réalisent. Ces paquets de traits morphosyntaxiques sont conçus comme des ensembles de traits structurés correspondant approximativement aux étiquettes morphosyntaxiques habituellement indiquées dans les gloses des formes. Chaque entrée lexicale est ainsi définie pour réaliser un ensemble de ces paquets de traits morphosyntaxiques.

Par ailleurs, chaque catégorie flexionnelle est elle-même définie pour réaliser un certain ensemble de paquets de traits morphosyntaxiques. Dans le cas canonique, l'ensemble de paquets de traits morphosyntaxiques réalisés par le paradigme d'une entrée lexicale équivaut donc à l'ensemble de paquets de traits morphosyntaxiques pour lequel la catégorie à laquelle il appartient est définie.

Les paquets de traits morphosyntaxiques peuvent être représentés comme des structures de traits.

Les cas de NEUTRALISATION peuvent ainsi être représentés dans  $\mathcal{PARSL}$  comme des cas où parmi les paquets de traits morphosyntaxiques réalisés par les formes d'une

entrée lexicale *i* donnée, il en existe au moins un qui subsume strictement<sup>1</sup> un paquet de traits morphosyntaxiques généralement réalisé par les formes des entrées appartenant canoniquement à la catégorie flexionnelle de *i*.

Les cas de SURDIFFÉRENCIATION sont les cas inverses : parmi les paquets de traits morphosyntaxiques réalisés par les formes d'une entrée lexicale *i* donnée, il en existe au moins un qui est une extension stricte<sup>2</sup> d'un paquet de traits morphosyntaxiques généralement réalisé par les formes des entrées appartenant canoniquement à la catégorie flexionnelle de *i*.

### C.3 Alignement des formes et des hiérarchies

Une autre façon, orthogonale, d'approcher la notion de hiérarchie de traits est de confronter les hiérarchies à des caractéristiques formelles affichées par les paradigmes.

Les données du slovaque que nous avons présentées dans le tableau 3.10 et appelées au tableau 4.3 présentent les caractéristiques suivantes :

- (a) Il s'agit de trois paradigmes nominaux, chacun affichant douze cases pour deux nombres et six cas.
- (b) Chacun des paradigmes est partitionnable en deux sous-paradigmes, en fonction de la valeur du trait nombre (un sous-paradigme du singulier et un sous-paradigme du pluriel).
- (c) Chacun des sous-paradigmes peut à son tour être partitionné de la même manière en six cases, une par cas.

Par ailleurs, les traits morphosyntaxiques exprimés par les paradigmes slovaques peuvent être représentés par un arbre illustrant la hiérarchie présentée à la figure C.1).

Selon ce partitionnement des paradigmes, le paradigme hétéroclite du nom OROL « *aigle* » emploie deux sous-paradigmes issus du premier partitionnement (b).

Dans la hiérarchie des traits exprimés par les noms slovaques, le nombre domine le cas, chaque nombre étant combinable avec chacun des six cas du slovaque. L'hétéroclise affichée par les noms d'animaux slovaques comme OROL « *aigle* » emploie des parties de paradigmes qui sont en accord avec cette hiérarchie : les sous-paradigmes concernés par l'hétéroclise comportent les formes servant à exprimer un sous-arbre de la hiérarchie de traits globale valable pour les noms du slovaque.

---

1. Les deux paquets de traits morphosyntaxiques ne sont donc pas identiques.  
2. Cf. la note ci-dessus.

FIGURE C.1 : Hiérarchie de traits possible pour les noms du slovaque

Ce découpage formel régulier qui revient pour tous les noms du slovaque peut être qualifié de *morphomique*.

## C.4 Morphomes et hiérarchies de traits

Dans la littérature actuelle, la notion de *morphome* peut avoir plusieurs acceptions. Les deux plus répandues sont :

1. Une fonction purement morphologique qui n'est aucunement conditionnée par des éléments extérieurs à la morphologie (la définition originelle d'Aronoff (1994)).
2. Une structure formelle morphologique régulière dans le système qui est en décalage par rapport à la hiérarchie de traits affichée par une langue donnée, cf. par exemple (Maiden, 2012 ; Blevins, 2006).

Le cas du slovaque évoqué ci-dessus correspond à une structure morphomique au sens de (1). En revanche, le découpage en sous-paradigmes concernés par l'hétéroclise apparaît comme motivé par la hiérarchie de traits de la figure C.1 et ne constitue donc pas un morphome au sens de (2). Un tel découpage des formes selon une hiérarchie de traits est ce que Corbett (2010) qualifie de FRACTURE FONCTIONNELLE (*feature split*) motivée.

Un autre exemple est celui du dhaasanac que nous avons présenté au tableau 2.1 du chapitre 2 et dont nous rappelons les données au tableau C.6.

Dans le paradigme du tableau C.6 les verbes SED « *marcher* » et YUUFUMI « *tousser* » présentent la même structure formelle, à savoir des syncrétismes entre 1.SG.EXCL, 3.SG.M, 1.PL.INCL, 3.PL.F et 3.PL.M d'une part, et 2.SG, 3.SG.F, 1.PL.EXCL et 2.PL de l'autre, sans qu'aucune caractéristique phonologique ou syntaxique valable en dhaasanac ne

	SED « <i>marcher</i> »		YUUFUMI « <i>tousser</i> »	
	SG	PL	SG	PL
1.INCL	—	<i>seḏ</i>	—	<i>yuufumi</i>
1.EXCL	<i>seḏ</i>	<i>sieti</i>	<i>yuufumi</i>	<i>yuufeeni</i>
2	<i>sieti</i>	<i>sieti</i>	<i>yuufeeni</i>	<i>yuufeeni</i>
3.F	<i>sieti</i>	<i>seḏ</i>	<i>yuufeeni</i>	<i>yuufumi</i>
3.M	<i>seḏ</i>	<i>seḏ</i>	<i>yuufumi</i>	<i>yuufumi</i>

TABLE C.6 : Morphome en dhaasanac (Corbett, 2010)

requière ces regroupements. Leur structure est ainsi morphomique au sens (1). Mais les regroupements syncrétiques observables ne se font pas selon une partition cohérente de la hiérarchie de traits morphosyntaxiques. Ils constituent également une structure morphomique au sens (2), autrement dit, un cas de fracture fonctionnelle non motivée.

Les données du slovaque comme les données du dhaasanac présentent ainsi des cas de morphomicité. On observe une fracture fonctionnelle régulière qui traverse la structure de tous les paradigmes du même type. La différence entre les données du slovaque et celles du dhaasanac réside dans la motivation de la fracture fonctionnelle.

Dans le cas du slovaque, la fracture suit un découpage prévu dans la hiérarchie de traits de la langue en question. La fracture est motivée par le rapport entre les traits habituellement observables. En d'autres termes, les traits observables concernés par la fracture fonctionnelle et intervenant dans la construction de la structure morphomique constituent un sous-arbre régulier de la hiérarchie de traits globale.

Dans le cas du dhaasanac aucune fracture motivée ne peut être observée par rapport à la hiérarchie globale. La fracture fonctionnelle apparaît comme non motivée, ce qui formellement se traduit par le fait qu'elle ne correspond pas à un sous-arbre de la hiérarchie globale.

Dans des cas de découpages morphomiques des paradigmes comme ceux du slovaque d'une part et du dhaasanac de l'autre, les fractures fonctionnelles motivées comme celles du slovaque seront considérées comme les plus canoniques selon Corbett (2010).

## C.5 Encodage de la morphomicité dans $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$

En *Network Morphology*, où grammaire, lexique et hiérarchies de traits sont conçus comme des hiérarchies d'héritage, les cas de fracture fonctionnelle non motivées *vs.* motivées sont aisément représentables. Les hiérarchies lexicales et grammaticales peuvent être comparées à une hiérarchie de traits indépendante. Une correspondance entre ces premières

et la hiérarchie de traits correspondrait ainsi à une fracture fonctionnelle motivée. Tout éloignement de cette correspondance suggérerait la présence d'une fracture fonctionnelle non motivée.

En  $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$ , ce type de correspondance directe ne fait pas partie des primitives du modèle. Toutefois, nous avons indiqué dans Walther (2012b) qu'une structuration externe des traits permet d'évaluer le caractère motivé ou non d'une structure fonctionnelle. Compte tenu de l'architecture du modèle, cette évaluation peut même se faire niveau par niveau. Cette évaluation dépasse cependant les primitives initialement conçues dans le modèle  $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$ . Il en va de même pour PFM.





## Transcription du sorani

Cette annexe comporte le système de transcription que nous avons utilisé pour rendre compte du système d'écriture du kurde sorani.

Le kurde sorani s'écrit de droite à gauche et emploie un système d'écriture alphabétique dérivé de l'alphabet arabo-persan qui est lui-même une version étendue de l'alphabet arabe. Il s'agit d'un alphabet comportant 36 caractères : 28 caractères arabes, 4 caractères ajoutés par le persan et 4 caractères spécifiques au kurde sorani.

Ces caractères et leur prononciation approchée sont indiqués dans le tableau D.1<sup>1</sup>. Nous avons adopté la transcription phonétique de Thackston (2006). Les caractères arabes d'origine y figurent sur fond blanc, les caractères ajoutés pour le persan sur fond gris clair et les caractères spécifiques au sorani sur un fond gris plus soutenu.

Bien que formellement présente dans l'alphabet officiel, certaines lettres ne sont guère utilisées en sorani, en dehors de leur emploi (facultatif) pour transcrire certains emprunts arabes. Nous avons indiqué ces caractères entre parenthèses dans le tableau D.1. Les ponctuations employées en sorani figurent dans le tableau D.2.

Une autre particularité de l'alphabet sorani concerne la transcription des voyelles. Tandis que l'alphabet arabo-persan ne transcrit alphabétiquement que les voyelles longues<sup>2</sup>, le kurde sorani a étendu l'alphabet pour également transcrire les voyelles brèves. Elles ont deux graphies distinctes selon qu'elles figurent en initiale ou en milieu de mot. En initiale, elles sont toujours précédées de la lettre *hamza* **ئ**. La seule voyelle à ne pas être transcrite en sorani est le [ɪ] court que nous avons transcrit « *i* » dans de cette thèse. Les voyelles du sorani et leur prononciation figurent dans les tableaux D.3 et D.4.

---

1. Les caractères sorani à gauche, les caractères en alphabet phonétique international (API) au milieu et la transcription à droite.

2. Les voyelles courtes peuvent être rendues par des diacritiques facultatifs.

ا	[a]	â	ب	[b]	b	ت	[t]	t	(ث)	[(s)]	s
ج	[dʒ]	j	ح	[h]	h	خ	[x]	kh	د	[d]	d
(ز)	[(z)]	(z)	ر	[r]	r	ړ	[r]	ř	ز	[z]	z
س	[s]	s	ش	[ʃ]	š	(ص)	[(s)]	(s)	(ض)	[(z)]	(z)
(ط)	[(t)]	(t)	(ظ)	[(z)]	(z)	ع	[ʔ]	'	غ	[ɣ]	gh
ف	[f]	f	ڤ	[v]/[w]	v	ق	[q]	q	ك	[l]	l
ل	[l]	l	ل	[ʎ]/[ɬ]	ɭ	م	[m]	m	ن	[n]	n
و	[æ]	a	و	[w]/[u]/[v]	u/w	ۆ	[o]	o	ی	[j]/[i]	y/î
پ	[p]	p	چ	[tʃ]	c	ژ	[ʒ]	ʒ	ک	[k]	k
گ	[g]	g	ه	[h]	h	ئ	[e]	ê			

TABLE D.1 : Alphabet kurde sorani

.	.	،	،	؛	؛	؟	؟
---	---	---	---	---	---	---	---

TABLE D.2 : Le système de ponctuation en kurde sorani

ئه	æ	ئى	e	ئى	i	ئو	u
ئا	a	ا	ɪ	ئو	o		

TABLE D.3 : Les voyelles initiales du kurde sorani

ه	æ	ئى	e	ى	i	و	u
ا	a			ۆ	o	وو	u

TABLE D.4 : Les voyelles du kurde sorani en milieu de mot

# Index



# Table des figures

0.1	Notation des éléments de la grammaire dans $\mathcal{PARSLI}$ . . . . .	xii
0.2	Notation pour une entrée lexicale standard dans $\mathcal{PARSLI}$ . . . . .	xiii
1.1	Le m-passif en tant que sous-paradigme des paradigmes de verbes transitifs	51
1.2	Définition des entrées lexicales pour les quatre types de conjugaison . . .	53
1.3	Définition de la morphologie verbale . . . . .	53
1.4	Définition des formes verbales . . . . .	53
1.5	Représentation de la déponence . . . . .	54
1.6	Règles de renvoi spécifiques pour les semi-déponents (verbes déponents mixtes) . . . . .	54
1.7	Entrées lexicales pour les déponents (verbes déponents) et semi-déponents (verbes déponents mixtes) . . . . .	55
1.8	Représentation sous forme de graphe de l'analyse de Hippiisley (2007) .	56
1.9	Modélisation des paradigmes des verbes déponents par syncrétisme . . .	57
1.10	Traits conjugationnels des radicaux et des désinences . . . . .	62
1.11	La passivation comme opération dérivationnelle . . . . .	63
1.12	Classes flexionnelles et désinences possibles . . . . .	64
1.13	Combinaison standard de zones pour les verbes actifs prototypiques . .	80
1.14	Combinaison standard de zones pour les verbes passifs prototypiques . .	81
2.1	Architecture de la grammaire selon le modèle de Halle (1973) . . . . .	115
2.2	Architecture de la grammaire chez Kiparsky (1982) . . . . .	118
2.3	Esquisse de l'architecture de la grammaire . . . . .	124
2.4	Représentation des super-classes en PFM . . . . .	132
2.5	Constructions des mots et <i>Correspondance Paradigmatique</i> . . . . .	133
2.6	Entrée lexicale du russe en <i>Network Morphology</i> . . . . .	138
2.7	Entrée lexicale du latin en <i>Network Morphology</i> . . . . .	139
2.8	Commande <code>#show DATR</code> . . . . .	140
2.9	Défauts en <i>Network Morphology</i> . . . . .	141
2.10	Entrée lexicale du polonais en <i>Network Morphology</i> . . . . .	142
3.1	Conception de standards en typologie . . . . .	163

3.2	Paradigmes canoniques . . . . .	172
3.3	Supplétion de radical . . . . .	173
3.4	Supplétion de forme . . . . .	174
3.5	Deux classes flexionnelles . . . . .	176
3.6	Paradigmes présentant des syncrétismes (lexème 1) et canonique (lexème 2)	177
3.7	Exemple d'hétéroclise . . . . .	178
3.8	Défectivité à la deuxième personne du pluriel (lexème 1) . . . . .	179
3.9	Paradigme surabondant (lexème 1) . . . . .	180
3.10	Axes de non-canonicté . . . . .	185
3.11	Approches en typologie . . . . .	188
4.1	Zones flexionnelles du présent de l'indicatif des verbes du premier groupe en français, c'est-à-dire de la classe flexionnelle $Z_{v1}^{exp}$ . . . . .	197
4.2	Sous-schémas (simplifiés) pour le BALAYER . . . . .	197
4.3	Classe et zones de radicaux (à gauche) et classe flexionnelle (à droite) du présent de l'indicatif des verbes du deuxième groupe en français . . . . .	202
4.4	Sous-schémas (simplifiés) pour le verbe <i>finir</i> . . . . .	202
4.5	Entrée lexicale pour le verbe ALLER . . . . .	203
4.6	Entrée lexicale pour le verbe FINIR . . . . .	203
4.7	Entrée lexicale pour le verbe ÊTRE . . . . .	204
4.8	Représentation des niveaux réalisationnels dans $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$ (exemple de l'analyse du latin donnée au chapitre 1) . . . . .	210
4.9	Entrée lexicale pour le nom du serbo-croate DUGME « bouton » . . . . .	217
4.10	Entrée lexicale pour le nom du serbo-croate ŽENA « femme » . . . . .	218
4.11	Entrée lexicale pour le nom serbo-croate DETE « enfant » . . . . .	218
4.12	Entrée lexicale pour le nom du serbo-croate JELEN « cerf » . . . . .	218
4.13	Entrée lexicale pour le nom du serbo-croate BRAT « frère » . . . . .	219
4.14	Entrée lexicale pour le verbe TRAIRE . . . . .	222
4.15	Entrée lexicale pour le verbe SEOIR . . . . .	223
4.16	Entrée lexicale pour le nom slovaque PRIATEL . . . . .	225
4.17	Axes de non-canonicté dans $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$ . . . . .	226
4.18	Notation des éléments de la grammaire dans $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$ . . . . .	229
4.19	Notation pour une entrée lexicale standard dans $\mathcal{P}\mathcal{A}\mathcal{R}\mathcal{S}\mathcal{L}$ . . . . .	231
4.20	Catégorie flexionnelle nominale en français . . . . .	234
4.21	Entrée lexicale pour le nom LIVRE . . . . .	234
4.22	Entrée lexicale pour le nom VIVRES . . . . .	234
4.23	Combinaison des espaces partitionnants inter-niveaux . . . . .	236
4.24	Entrée lexicale pour le verbe BALAYER . . . . .	237
4.25	Entrée lexicale pour le verbe latin AGGREDIOR . . . . .	239
4.26	Entrée lexicale pour le nom serbo-croate DETE « enfant » . . . . .	243
4.27	Entrée lexicale pour le nom serbo-croate BRAT « frère » . . . . .	244
4.28	Entrée lexicale défective : le verbe TRAIRE . . . . .	251
4.29	Entrée lexicale surabondante : le verbe ASSEOIR . . . . .	253
4.30	Entrée lexicale pour le nom slovaque OROL . . . . .	261

4.31	Entrée lexicale pour le verbe ALLER . . . . .	263
4.32	Entrée lexicale pour le verbe ÊTRE . . . . .	264
4.33	Entrée lexicale pour le verbe SEOIR . . . . .	268
4.34	Entrée lexicale pour le nom slovaque PRIATEL . . . . .	270
4.35	Entrée lexicale pour le nom serbo-croate DETE « <i>enfant</i> » . . . . .	271
5.1	Entrées lexicales verbales intensionnelles extraites du <i>Lefff</i> . . . . .	285
5.2	Entrées lexicales verbales extensionnelles produites à partir du <i>Lefff</i> . . . . .	286
5.3	Entrées lexicales adjectivales intensionnelles extraites du <i>Lefff</i> . . . . .	286
5.4	Entrées lexicales adjectivales extensionnelles produites à partir du <i>Lefff</i> . . . . .	287
5.5	Classes de graphèmes/phonèmes en Alexina (données du <i>Lefff</i> original)	287
5.6	Règles morphographémiques/morphologiques en Alexina (données du <i>Lefff</i> original) . . . . .	288
5.7	La classe flexionnelle $adj-1$ dans la description morphologique du <i>Lefff</i> original . . . . .	288
5.8	La classe flexionnelle $nc-4$ dans la description morphologique du <i>Lefff</i> original . . . . .	289
5.9	La classe flexionnelle $adj-4$ dans la description morphologique du <i>Lefff</i> original . . . . .	289
5.10	Exemple d'une règle de renvoi intra-classe en Alexina extraite de PolLex . . . . .	290
5.11	Règles de réalisations incompatibles et variantes en Alexina ( <i>Lefff</i> original)	290
5.12	Allomorphie radicale régulière en Alexina <sub>PARSLI</sub> (données de MaltLex) . . . . .	294
5.13	Supplétion de radical en Alexina (données du <i>Lefff</i> modifié) . . . . .	295
5.14	La définition des niveaux réalisationnels dans Alexina <sub>PARSLI</sub> (données de MaltLex) . . . . .	296
5.15	La supplétion de formes dans Alexina <sub>PARSLI</sub> (données du <i>Lefff</i> modifié) . . . . .	297
5.16	La supplétion surabondante de formes dans Alexina <sub>PARSLI</sub> : entrée intensionnelle standard (données du <i>Lefff</i> d'origine) . . . . .	297
5.17	La supplétion surabondante de formes dans Alexina <sub>PARSLI</sub> : forme supplétive supplémentaire (données du <i>Lefff</i> original) . . . . .	297
5.18	La définition (directe) des zones dans Alexina <sub>PARSLI</sub> (données de <i>Leffa</i> modifié) . . . . .	298
5.19	La définition des zones dans les schèmes d'Alexina <sub>PARSLI</sub> (données de <i>Leffa</i> )	298
5.20	La définition de classes réalisationnelles par référence à des zones indépendantes dans Alexina <sub>PARSLI</sub> (données de <i>Leffa</i> ) . . . . .	300
5.21	La définition des espaces partitionnants dans Alexina <sub>PARSLI</sub> au niveau radical (données de MaltLex) . . . . .	301
5.22	La définition des espaces partitionnants dans Alexina <sub>PARSLI</sub> au niveau exposant (données de <i>Leffa</i> ) . . . . .	301
5.23	La définition des zones par les espaces partitionnants dans Alexina <sub>PARSLI</sub> (données de <i>Leffa</i> ) . . . . .	302
5.24	La représentation des schèmes dans les entrées intensionnelles d'Alexina <sub>PARSLI</sub> (données de <i>Leffa</i> ) . . . . .	303



5.25	La représentation des schèmes dans les entrées intensionnelles d'Alexina <sub>PARSLI</sub> (données de <i>Leffla</i> ) . . . . .	304
5.26	Entrées lexicales intensionnelles pour les noms du serbo-croate en Alexina <sub>PARSLI</sub> . . . . .	305
5.27	La représentation des fonctions de transfert dans les schèmes d'Alexina <sub>PARSLI</sub>	306
5.28	Classes de graphèmes/phonèmes Alexina <sub>PARSLI</sub> (données de MaltLex) . .	307
5.29	Opérations morphologiques en Alexina <sub>PARSLI</sub> (données de MaltLex) . .	308
5.30	Opérations morphologiques avec argument(s) en Alexina <sub>PARSLI</sub> (MaltLex)	308
5.31	Opérations morphologiques avec argument(s) en Alexina <sub>PARSLI</sub> ( <i>Leffla</i> ) .	309
5.32	La définition des cases à réaliser dans Alexina <sub>PARSLI</sub> (données de MaltLex)	310
5.33	Définition de la défektivité (données de <i>Leffla</i> ) . . . . .	310
5.34	Définition de la défektivité (données de <i>Leffla</i> ) . . . . .	311
5.35	La représentation des règles de réalisation dans Alexina <sub>PARSLI</sub> (données de MaltLex) . . . . .	312
5.36	Représentation des CSR . . . . .	323
5.37	Opérations morphologiques sur les radicaux définies (données de MaltLex)	332
5.38	Règles de formation de radicaux pour les verbes à base verbale CVCVC (données de MaltLex) . . . . .	334
5.39	Règle de réalisation standard de la voyelle thématique CVCVC (données de MaltLex) . . . . .	335
5.40	Règles de réalisation spécifiques de la voyelle thématique CVCVC (données de MaltLex) . . . . .	335
5.41	Classes de caractères définis pour le maltais (données de MaltLex) . . . .	335
5.42	Règles morphologiques liées à la voyelle thématique (données de MaltLex) . . . . .	336
5.43	Zone athématique (données de MaltLex) . . . . .	336
5.44	Classe flexionnelle unique des verbes maltais (données de MaltLex) . . .	336
5.45	Entrées intensionnelles (données de MaltLex) . . . . .	337
5.46	Schème des verbes non surabondants (données de MaltLex) . . . . .	340
5.47	Schème de SATAR « <i>protéger</i> » (données de MaltLex) . . . . .	341
5.48	Schème de QALEB « <i>changer</i> » (données de MaltLex) . . . . .	341
5.49	Entrées extensionnelles de QALEB « <i>changer</i> » (données de MaltLex) . .	342
6.1	Esquisse de l'architecture de la grammaire . . . . .	348
6.2	Zones flexionnelles du présent de l'indicatif des verbes du premier groupe en français . . . . .	355
6.3	sous-schèmes (simplifiés) pour le verbe <i>balayer</i> . . . . .	356
6.4	Classe et zones de radicaux (à gauche) et classe flexionnelle (à droite) du présent de l'indicatif des verbes du deuxième groupe en français . . . . .	357
6.5	Sous-schèmes (simplifiés) pour le verbe <i>finir</i> . . . . .	357
6.6	Caractéristiques principales des quatre descriptions de la morphologie verbale du français . . . . .	361
6.7	Longueurs de description de différentes descriptions de la morphologie verbale du français . . . . .	362

6.8	Poids des syllabes maltaises . . . . .	371
6.9	Ordre préférentiel d'effacement des voyelles surnuméraires dans l'ordre suivant : $-1, -2 \rightarrow -n, 1, 2 \rightarrow n$ . . . . .	371
6.10	Hiérarchie des contraintes et règles (mor-)phonologiques . . . . .	376
6.11	Perfectif des verbes à base verbale CVCVC . . . . .	379
6.12	Imperfectif des verbes à base verbale CVCVC . . . . .	380
6.13	RAD6 alternatif pour les verbes à base verbale CVRVC . . . . .	380
6.14	Perfectif des verbes à base verbale CVCCVC . . . . .	382
6.15	Imperfectif des verbes à base verbale CVCCVC . . . . .	382
6.16	Perfectif des verbes à base verbale CVCV . . . . .	383
6.17	Imperfectif des verbes à base verbale CVCV . . . . .	384
6.18	Perfectif des verbes à base verbale CVCC . . . . .	385
6.19	Imperfectif des verbes à base verbale CVCC . . . . .	386
6.20	Perfectif des verbes à base verbale CV:C . . . . .	387
6.21	Imperfectif des verbes à base verbale CV:C . . . . .	388
6.22	Représentation des mesures de compacité sur les descriptions du maltais	390
7.1	Catégorie flexionnelle verbale en latin . . . . .	417
7.2	Analyse 3H : entrée lexicale du flexème CONTINEO « <i>contenir</i> » . . . . .	417
7.3	Analyse 3H : entrée lexicale du flexème AUDIOR « <i>être entendu</i> » . . . . .	418
7.4	Analyse 3H : entrée lexicale du flexème ACCIPIO « <i>accepter</i> » . . . . .	418
7.5	Analyse 3H : entrée lexicale du flexème ACCIPIOR « <i>être accepté</i> » . . . . .	419
7.6	Classes réalisationnelles dans l'analyse 3H . . . . .	424
7.7	Analyse 3H : entrée lexicale du flexème de la base AMO « <i>aimer</i> » . . . . .	425
7.8	Analyse 3H : entrée lexicale du flexème AMOR « <i>être aimé</i> » dérivé de AMO « <i>aimer</i> » . . . . .	426
7.9	Analyse 3H : entrée lexicale du flexème ACCIPIO « <i>accepter</i> » . . . . .	426
7.10	Analyse 3H : entrée lexicale du flexème ACCIPIOR « <i>être accepté</i> » . . . . .	427
7.11	Analyse 3H : entrée lexicale du flexème OPINOR « <i>penser</i> » . . . . .	427
7.12	Analyse 3H : entrée lexicale du flexème LOQUOR « <i>parler</i> » . . . . .	428
7.13	Analyse 4H : entrée lexicale du flexème CONTINEO « <i>contenir</i> » . . . . .	431
7.14	Analyse 4H : entrée lexicale du flexème CONTINEOR « <i>être contenu</i> » . . . . .	431
7.15	Analyse 4H : entrée lexicale du flexème AUDIOR « <i>être entendu</i> » . . . . .	431
7.16	Analyse 4H : entrée lexicale du flexème ACCIPIO « <i>accepter</i> » . . . . .	432
7.17	Analyse 4H : entrée lexicale du flexème ACCIPIOR « <i>être accepté</i> » . . . . .	432
7.18	Classes réalisationnelles dans l'analyse 4H . . . . .	433
7.19	Analyse 4H : entrée lexicale du flexème OPINOR « <i>penser</i> » . . . . .	434
7.20	Analyse 4H : entrée lexicale du flexème LOQUOR « <i>parler</i> » . . . . .	434
7.21	Mesures de compacité globale (lexique + grammaire)s des deux descriptions de la flexion verbale latine . . . . .	436
7.22	Mesures de compacité pour la seule grammaire flexionnelle des deux descriptions de la flexion verbale latine . . . . .	438
7.23	Analyse 3H : Entrée lexicale du flexème de la base GERO « <i>tourner</i> » sans hétéroclise au niveau radical . . . . .	441

---

7.24	Analyse 3H : Entrée lexicale du flexème de la base GERO « <i>tourner</i> » avec hétéroclise au niveau radical . . . . .	441
7.25	3F : catégorie flexionnelle verbale en latin . . . . .	444
7.26	Analyse 3F : entrée lexicale du flexème AMO « <i>aimer</i> » . . . . .	444
7.27	Analyse 3F : entrée lexicale du flexème LUGEO « <i>briller</i> » . . . . .	445
7.28	Analyse 3F : entrée lexicale du flexème LOQUOR « <i>briller</i> » . . . . .	446
7.29	Mesures de compacité globale (lexique + grammaire)s des trois descriptions de la flexion verbale latine . . . . .	447
7.30	Analyse 4H : entrée lexicale du flexème COEPI « <i>avoir commencé</i> » . . . . .	453
7.31	Mesures de canonicité de nos descriptions de la flexion verbale latine . . . . .	471
7.32	Attribution des classes flexionnelles pour les flexèmes des verbes m-actifs dérivés dans l'analyse 3H . . . . .	472
7.33	Attribution des classes flexionnelles pour les flexèmes des verbes m-actifs dérivés dans l'analyse 4H . . . . .	472
A.1	Représentation des super-classes verbales correspondant à l'analyse du chapitre 1 . . . . .	490
C.1	Hiérarchie de traits possible pour les noms du slovaque . . . . .	507

# Liste des tableaux

0.1	Flexion verbale du népali (données de Bonami & Boyé (2010b)) . . . . .	14
1.1	Formes de citations avec indication des trois radicaux . . . . .	33
1.2	Voyelle thématique <i>a</i> — exemple AMO « <i>aimer</i> » . . . . .	35
1.3	Voyelle thématique <i>e</i> — exemple CONTINEO « <i>contenir</i> » . . . . .	36
1.4	Voyelle thématique <i>i</i> — exemple AUDIO « <i>entendre</i> » . . . . .	37
1.5	Verbes athématiques — exemple DUCO « <i>conduire, mener</i> » . . . . .	38
1.6	Verbes à thème mixte — exemple ACCIPIO « <i>accepter</i> » . . . . .	39
1.7	Répartition des types de verbes en latin (données de Prior & Wohlberg (1995)) . . . . .	41
1.8	Désinences actives et passives au présent de l'indicatif imperfectif . . . . .	41
1.9	Alignements possibles entre la morphologie verbale et la syntaxe pour les verbes latins . . . . .	42
1.10	Les désinences m-actives . . . . .	45
1.11	Les désinences m-passives . . . . .	46
1.12	Constructions valentielles disponibles avec un marquage m-actif/m-passif selon Flobert (1967) . . . . .	50
1.13	Synchrétismes dans le paradigme d'un nom neutre en latin (BELLUM « <i>la guerre</i> ») . . . . .	60
1.14	Nombre de verbes concernés par les valeurs alternatives du m-passif selon Flobert (1967) . . . . .	66
1.15	Formes de la classe flexionnelle A (verbes actifs canoniques) . . . . .	69
1.16	Formes de la classe flexionnelle B (verbes passifs canoniques) . . . . .	70
1.17	Zones flexionnelles de la classe flexionnelle A . . . . .	74
1.18	Zones flexionnelles de la classe flexionnelle B . . . . .	75
1.19	Zones thématiques de la classe flexionnelle A . . . . .	78
1.20	Zones thématiques de la classe flexionnelle B . . . . .	79
1.21	Zones flexionnelles de la classe flexionnelle C . . . . .	83
1.22	Zones flexionnelles de la classe flexionnelle D . . . . .	84
1.23	Répartition des types de verbes dans <i>Leffla</i> . . . . .	85
1.24	Inventaire des traits morphosyntaxiques intervenant dans l'analyse . . . . .	87
1.25	Classes de position pour les verbes de type A . . . . .	88

1.26	Classes de position pour les verbes de type B . . . . .	88
1.27	Classes de position pour les verbes de type C . . . . .	89
1.28	Classes de position pour les verbes de type D . . . . .	89
1.29	Informations lexicales associées aux verbes latins . . . . .	92
1.30	Les quatre classes flexionnelles du latin . . . . .	94
2.1	Morphème en dhaasanac (Corbett, 2010) (données de Tosco (2001)) . .	102
2.2	Radical et exposants dans le paradigme du verbe MANGER au présent de l'indicatif . . . . .	112
2.3	Allomorphie du radical dans le paradigme du verbe <i>boire</i> au présent de l'indicatif . . . . .	113
2.4	Allomorphie du radical dans le paradigme du verbe <i>boire</i> selon Bonami & Boyé (2003) . . . . .	113
2.5	Le paradigme du verbe <i>boire</i> sans alternance de radicaux . . . . .	114
2.6	Typologie des approches formelles de la morphologie (Stump, 2001) . .	126
2.7	Classes de position des verbes sorani . . . . .	129
2.8	Nominatifs et accusatifs de trois noms latins . . . . .	130
3.1	Inventaire de cas de réduplication d'après Rubino (2011) . . . . .	152
3.2	Les types d'universaux chez Greenberg (1963) . . . . .	154
3.3	Différences entre flexion et dérivation canoniques selon Brown & Hippisley (2012, p. 37) . . . . .	168
3.4	Critères pour la flexion canonique selon Corbett (2007a). . . . .	169
3.5	Critères supplémentaires pour la flexion canonique . . . . .	170
3.6	Supplétion de formes dans le paradigme du verbe <i>être</i> . . . . .	174
3.7	Radicaux du présent (RAD1) et du passé (RAD2) en persan . . . . .	174
3.8	Radicaux irréguliers du passif en kurde sorani . . . . .	175
3.9	Classes flexionnelles en latin . . . . .	176
3.10	Hétéroclise dans les paradigmes des noms d'animaux slovaques . . . . .	178
3.11	Exemple de paradigme défectif slovaque : VIANOCE . . . . .	180
3.12	Surabondance en italien d'après Thornton (2010) . . . . .	181
3.13	Surabondance en français pour le verbe ASSEoir . . . . .	181
3.14	Surabondance à l'indicatif présent en français pour le verbe BALAYER . .	182
3.15	Déclinaison régulière de noms du serbo-croate . . . . .	183
3.16	Décalage dans deux paradigmes de noms du serbo-croate . . . . .	184
3.17	Récapitulatif des huit types de phénomènes non canoniques . . . . .	186
4.1	Surabondance à l'indicatif présent en français pour le verbe BALAYER . .	195
4.2	Représentation polyparadigmatique des verbes en <i>-ayer</i> (exemple de l'indicatif présent de BALAYER) . . . . .	196
4.3	Hétéroclise dans les paradigmes des noms d'animaux slovaques . . . . .	199
4.4	Allomorphie radicale régulière (exemple des radicaux du présent (RAD1) et du passé (RAD2) en persan) . . . . .	200
4.5	Supplétion de radicaux dans le paradigme du verbe ALLER . . . . .	200

4.6	Voyelle thématique <i>i</i> latins — exemples AUDIO « <i>entendre</i> » et AUDIOR « <i>être entendu</i> » . . . . .	206
4.7	Verbes athématiques latins — exemples DUCO « <i>conduire, mener</i> » et DUCOR « <i>être conduit</i> » . . . . .	207
4.8	Verbes latins à thème mixte — exemples ACCIPIO « <i>accepter</i> » et ACCIPIOR « <i>être accepté</i> » . . . . .	208
4.9	Exponence multiple en chinantec de Tlatepuzco (données de Palancar (à paraître)) . . . . .	211
4.10	Déclinaison régulière de noms du serbo-croate . . . . .	216
4.11	Décalage dans deux paradigmes de noms du serbo-croate . . . . .	216
4.12	Paradigme déficient : le verbe SEOIR . . . . .	222
4.13	Paradigmes masculins animés consonantiques en slovaque, dont l'un manifeste de la surdifférenciation au VOC.SG (données de (Zauner, 1973)).	224
4.14	Récapitulatif des neuf types de phénomènes non canoniques dans פֿאַרשלי	227
4.15	Zones de radicaux dans le paradigme du verbe AGGREDIOR « <i>attaquer</i> »	240
4.16	Zones thématiques dans le paradigme du verbe AGGREDIOR « <i>attaquer</i> »	241
4.17	Zones flexionnelles dans le paradigme du verbe AGGREDIOR « <i>attaquer</i> »	242
4.18	Surabondance en français pour le verbe ASSEOIR . . . . .	253
4.19	Synchrétismes dans les paradigmes de noms latins . . . . .	266
4.20	Ensembles de cases syncrétiques dans les paradigmes de noms latins . . .	266
5.1	Exemples de FMLS de quelques adjectifs du français (Bonami & Boyé, 2005)	281
5.2	Sous-paradigme du singulier de quelques adjectifs du français (Bonami & Boyé, 2005) . . . . .	281
5.3	Les lexiques Alexina . . . . .	283
5.4	Paradigmes des verbes RASS et MESS en maltais . . . . .	293
5.5	Déclinaison régulière de noms en serbo-croate . . . . .	305
5.6	Décalage dans deux paradigmes de noms en serbo-croate . . . . .	305
5.7	Marques affixales portées par les verbes du premier binyan . . . . .	315
5.8	Paradigmes de verbes du premier binyan . . . . .	316
5.9	Différenciation maximale et répartition des radicaux dans un paradigme maltais . . . . .	316
5.10	Cases à radicaux systématiquement syncrétiques . . . . .	319
5.11	Structures des radicaux dans les paradigmes du premier binyan . . . . .	321
5.12	Notation de la différenciation maximale dans les classes de syncrétisme radical (CSR) dans (Camilleri, 2011) . . . . .	322
5.13	Changements vocaliques pour les verbes à base CVCVC . . . . .	325
5.14	Changements vocaliques pour les verbes à base CVCC . . . . .	325
5.15	Changements vocaliques pour les verbes à base CVC . . . . .	326
5.16	Changements vocaliques pour les verbes à base CVCV . . . . .	326
5.17	Changements vocaliques pour les verbes à base CVCCVC . . . . .	327
5.18	Changements vocaliques dans les CSR du premier binyan . . . . .	328
5.19	Paradigmes surabondants en maltais . . . . .	330
5.20	Les trois niveaux réalisationnels l'analyse פֿאַרשלי du maltais . . . . .	333

6.1	Entrées lexicales pour quelques flexèmes dans chacune de nos quatre représentations concurrentes de la flexion verbale du français . . . . .	359
6.2	Paradigmes de verbes du premier binyan . . . . .	366
6.3	Structures des radicaux dans les paradigmes du premier binyan . . . . .	367
6.4	Correspondance entre graphèmes et phonèmes en maltais . . . . .	368
6.5	Paradigmes de verbes CVCVC . . . . .	378
6.6	Paradigmes de verbes à base verbale CVCCVC . . . . .	381
6.7	Paradigmes de verbes à base verbale CVCV . . . . .	383
6.8	Paradigmes de verbes à base verbale CVCC . . . . .	384
6.9	Paradigmes de verbes à base CV:C . . . . .	387
6.10	Résultats des mesures de compacité sur les descriptions du maltais . . . . .	389
6.11	Classes de positions affixales dans une analyse PFM du système verbal sorani	393
6.12	Marquage de la personne en sorani . . . . .	394
6.13	Positions possibles pour les MPM à l'intérieur d'un verbe transitif . . . . .	395
7.1	Synchrétismes dans les paradigmes des verbes à thème en <i>a</i> — exemples d'AMO « <i>aimer</i> » et AMOR « <i>être aimé</i> » . . . . .	411
7.2	Synchrétismes dans les paradigmes des verbes à thème en <i>e</i> — exemples de CONTINEO « <i>contenir</i> » et CONTINEOR « <i>être contenu</i> » . . . . .	412
7.3	Synchrétismes dans les paradigmes des verbes à thème en <i>i</i> — exemples d'AUDIO « <i>entendre</i> » et AUDIOR « <i>être entendu</i> » . . . . .	413
7.4	Synchrétismes dans les paradigmes des verbes athématiques — exemples de DUCO « <i>conduire, mener</i> » et DUCOR « <i>être conduit, mené</i> » . . . . .	414
7.5	Synchrétismes dans les paradigmes des verbes à thème mixte — exemples d'ACCIPIO « <i>accepter</i> » et ACCIPIOR « <i>être accepté</i> » . . . . .	415
7.6	Classes thématiques définies au niveau 2 de l'analyse 3H . . . . .	420
7.7	Classes d'exponence définies au niveau 2 de l'analyse 3H . . . . .	421
7.8	Classes de radicaux définies . . . . .	423
7.9	Classes d'exponence définies au niveau 3 et 4 de l'analyse 4H . . . . .	430
7.10	Résultats des mesures de compacité sur les deux descriptions de la flexion verbale latine (en bits) . . . . .	437
7.11	Résultats des mesures de compacité sur les trois descriptions de la flexion verbale latine (en bits) . . . . .	448
7.12	Types de déficience de la flexion verbale latine selon 3H/4H . . . . .	455
7.13	Types de déficience de la flexion verbale latine selon 3F . . . . .	456
7.14	Récapitulatif des mesures de canonicité pour le système flexionnel verbal latin . . . . .	470
C.1	Le paradigme de l'adjectif BRUN . . . . .	502
C.2	Le paradigme de l'adjectif PRATIQUE . . . . .	502
C.3	Le paradigme de l'adjectif GRIS . . . . .	502
C.4	Le paradigme verbal KITEB « <i>lire</i> » (maltais) . . . . .	503
C.5	Le verbe polonais NIEŚĆ « <i>porter</i> » au présent et au prétérit (Grappin, 1985)	504
C.6	Morphome en dhaasanac (Corbett, 2010) . . . . .	508

D.1	Alphabet kurde sorani . . . . .	512
D.2	Le système de ponctuation en kurde sorani . . . . .	512
D.3	Les voyelles initiales du kurde sorani . . . . .	512
D.4	Les voyelles du kurde sorani en milieu de mot . . . . .	512





# Bibliographie



# Bibliographie

- Farrell ACKERMAN, James P. BLEVINS, Robert MALOUF (2009). “Parts and wholes : Implicative patterns in inflectional paradigms”. In : J. P. BLEVINS, J. BLEVINS (éds.), *Analogy in Grammar : Form and Acquisition*, Oxford University Press, p. 54–81.
- Farrell ACKERMAN, Gregory T. STUMP (2004). “Paradigms and periphrastic expression : A study in realization-based lexicalism”. In : Andrew SPENCER, Louisa SADLER (éds.), *Projecting Morphology*, Stanford, Californie, États Unis d’Amérique : CSLI, pp. 111–157.
- Stephen R. ANDERSON (1982). “Where’s Morphology ?” In : *Linguistic Inquiry* 13, pp. 571–612.
- (1992). *A-morphous Morphology*. Cambridge, Royaume Uni : Cambridge University Press.
- (2005). *Aspects of the theory of clitics*. Oxford, Royaume Uni : Oxford University press.
- François ANDRY, Norman M. FRASER, Scott McGLASHAN, Nick J. YOUD (1993). “Making DATR work for speech : lexicon compilation in SUNDIAL”. In : *Computational Linguistics* 18, 3, pp. 245–267. Special issue on inheritance : II.
- Haruo AOKI (1963). “Reduplication in Nez Perce”. In : *International Journal of American Linguistics* 29, pp. 42–44.
- M. ARONOFF (1994). *Morphology by Itself: Stems and Inflectional Classes*. Linguistic Inquiry Monographs. MIT Press.
- Michel ARRIVÉ (éd.) (1997). *Bescherelle : La conjugaison pour tous*. Paris : Didier Hatier.
- Eric Ormerod ASHTON (1947). *Swahili Grammar*. London : Longmans, Green & Co. 2nd edition.
- William BADECKER, Alfonzo CARAMAZZA (1989). “A lexical distinction between inflection and derivation”. In : *Linguistic Inquiry* 20, pp. 108–116.

- Matthew BAERMAN (2006). “Deponency in Serbo-Croatian”. Base de données en ligne. Typological Database on Deponency, Surrey Morphology Group, CMC, University of Surrey. Consulté en été 2012, URL <http://www.smg.surrey.ac.uk/deponency/Examples/Serbo-Croatian.htm>.
- (2007a). “Morphological reversals”. In : *Journal of Linguistics* 43, p. 33–61.
- (2007b). “Morphological Typology of Deponency”. In : Matthew BAERMAN, Greville G. CORBETT, Dunstan BROWN, Andrew HIPPISEY (éds.), *Deponency and Morphological Mismatches*. The British Academy, Oxford, Royaume-Uni : Oxford University Press, t. 145, pp. 1–19.
- Matthew BAERMAN, Greville G. CORBETT (2012). “Stem alternations and multiple exponence”. In : *Word Structure* 5, 1, pp. 52–68.
- Max BANE (2008). “Quantifying and Measuring Morphological Complexity”. In : Charles B. CHANG, Hannah J. HAYNIE (éds.), *Proceedings of the 26th West Coast Conference on Formal Linguistics*. Sommerville, États-Unis.
- Petra BARG (1994). “Automatic Acquisition of DATR Theories from Observations”. In : *Theorie des Lexikons : Arbeiten des Sonderforschungsbereichs 282*, t. 59.
- Clive H. BEAUMONT (1979). *The Tigak Language of New Ireland*, t. 58 de *Pacific Linguistics, Series B*. Canberra : Australian National University.
- Kenneth R. BEESLEY, Lauri KARTTUNEN (2003). *Finite State Morphology*. Studies in Computational Linguistics. CSLI Publications.
- Charles Edwin BENNETT (1907). *The Latin Language : A Historical Outline of the Sounds, Inflections, and Syntax*. Bennett’s Latin series. Allyn and Bacon.
- Émile BENVENISTE (1966). *Problèmes de linguistique générale*. Paris : Gallimard.
- Helma VAN DEN BERG (1995). *A Grammar of Hunzib*, t. 1 de *Lincom Studies in Caucasian Linguistics*. München : Lincom Europa.
- Balthasar BICKEL, Johanna NICHOLS (2002). “Autotypologizing databases and their use in fieldwork”. In : *Int. LREC Workshop on Resources and Tools in Field Linguistics*. Las Palmas de Gran Canaria, Espagne.
- James P. BLEVINS (2003). “Passives and impersonals”. In : *Journal of Linguistics* 39, pp. 473–520.
- (2006). “Word-based morphology”. In : *Journal of Linguistics* pp. 531–573.
- (2012). “Word-based morphology from Aristotle to modern WP”. In : K. ALLAN (éd.), *Oxford Handbook of the History of Linguistics*, Oxford University Press, p. 396–417.

- Leonard BLOOMFIELD (1933). *Language*. New-York, NY, États Unis d'Amérique : University of Chicago Press.
- (1939). “Menomini morphophonemics”. In : *Études phonologiques dédiées à la mémoire de M. le prince N.S. Trubetzkoy*, Prague, République Tchèque, t. 8 de *Travaux du Cercle Linguistique de Prague*, pp. 105–115. Réimprimé dans Hockett (1970), pages 351-362.
- Olivier BONAMI (2012). “Introduction”. In : *Word Structure* 5, 1, pp. 1–6.
- Olivier BONAMI, Gilles BOYÉ (2010a). “La morphologie flexionnelle est-elle une fonction ?” In : I. CHOI-JONIN, M. DUVAL, O. SOUTET (éds.), *Typologie et comparatisme. Hommages offerts à Alain Lemaréchal*, Louvain, Belgique : Peeters, pp. 21–35.
- (2010b). “Opaque paradigms, transparent forms in Nepali conjugation”. Communication orale au Workshop On Theoretical Morphology 5.
- Olivier BONAMI, Gilles BOYÉ, Hélène GIRAUDO, Madeleine VOGA (2008). “Quels verbes sont réguliers en français ?” In : *Actes du premier Congrès Mondial de Linguistique Française*. pp. 1511–1523.
- Olivier BONAMI, Gilles BOYÉ (2002). “Suppletion and dependency in inflectional morphology”. In : F. Van EYNDE, L. HELLAN, D. BEERMAN (éds.), *The Proceedings of the HPSG '01 Conference*. Stanford, États Unis d'Amérique : CSLI Publications.
- (2003). “Supplétion et classes flexionnelles dans la conjugaison du français”. In : *Langages* 152, pp. 102–126.
- (2005). “Construire le paradigme d'un adjectif”. In : *Recherches Linguistiques de Vincennes* pp. 77–98.
- (2006). “Deriving inflectional irregularity”. In : *Proceedings of the 13th International Conference on HPSG*. Stanford, États Unis d'Amérique : CSLI Publications, pp. 39–59.
- Olivier BONAMI, Pollet SAMVELIAN (2008). “Sorani Kurdish person markers and the typology of agreement”. Communication orale au 13th International Morphology Meeting.
- Olivier BONAMI, Gregory T. STUMP (). “Paradigm Function Morphology”. In : Andrew SPENCER (éd.), *Handbook of Morphology*, Blackwell. 2 éd. à paraître.
- Geert. E. BOOIJ (1993). “Against split morphology”. In : Geert BOOIJ, Jaap VAN MARLE (éds.), *Yearbook of Morphology 1993*, Boston, États Unis d'Amérique/Dordrecht, Pays-Bas : Kluwer, pp. 27–49.
- (1996). “Inherent versus contextual inflection and the split morphology hypothesis.” In : Geert BOOIJ, Jaap VAN MARLE (éds.), *Yearbook of Morphology 1995*, Boston, États Unis d'Amérique/Dordrecht, Pays-Bas : Kluwer, pp. 1–16.

- (2010). *Construction Morphology*. Oxford linguistics. Oxford University Press.
- J.-P. BOONS, A. GUILLET, C. LECLÈRE (1976). “La structure des phrases simples en français, Classes de constructions transitives”. Rap. tech., LADL, CNRS, Paris 7.
- Gilles BOYÉ (2000). *Problèmes de morpho-phonologie verbale en français, espagnol et italien*. Thèse de doctorat, Université Paris 7.
- (2011). “Régularité et classes flexionnelles dans la conjugaison du français”. In : Michel ROCHÉ, Gilles BOYÉ, Nabil HATHOUT, Stéphanie LIGNON, Marc PLÉNAT (éds.), *Des unités morphologiques au lexique*, Hermes Science.
- Gilles BOYÉ (2006). “Suppletion”. In : Keith BROWN (éd.), *Encyclopedia of Language and Linguistics (2nd ed.)*, Oxford, Royaume Uni : Elsevier, t. 12, pp. 297–299.
- Gilles BOYÉ, Patricia CABREDO HOFHERR (2010). “Defectiveness as Stem Suppletion in French and Spanish Verbs”. In : Matthew BAERMAN, Greville G. CORBETT, Dunstan BROWN (éds.), *Defective Paradigms : Missing forms and what they tell us*, Oxford, Royaume Uni : The British Academy, pp. 35–52.
- Joan BRESNAN (éd.) (1982). *The Mental Representation of Grammatical Relations*. MIT Press.
- Joan BRESNAN (2001). *Lexical-Functional Syntax*. Oxford, Royaume-Uni : Blackwell.
- Joan BRESNAN, Sam A. MCHOMBO (1995). “The Lexical Integrity Principle : evidence from Bantu”. In : *Natural Language and Language Theory 13*, p. 181–254.
- Dunstan BROWN, Marina CHUMAKINA, Greville G. CORBETT, Gergana POPOVA, Andrew SPENCER (2012). “Defining ‘periphrasis’ : key notions”. In : *Morphology 22*, 2.
- Dunstan BROWN, Greville G. CORBETT, Norman FRASER, Andrew HIPPISEY, Alan TIMBERLAKE (1996). “Russian noun stress and network morphology”. In : *Linguistics 34*, pp. 53–107.
- Dunstan BROWN, Roger EVANS (2012). “Inflectional defaults and principal parts : an empirical investigation”. In : *Proceedings of the 17th International Conference on HPSG*. Stanford, États Unis d’Amérique : CSLI Publications, pp. 235–254.
- Dunstan BROWN, Andrew HIPPISEY (2012). *Network Morphology : A Defaults-based Theory of Word Structure*. Cambridge University Press.
- Luigi BURZIO (2002). “Missing Players : Phonology and the Past-tense Debate”. In : *Lingua 112*, pp. 157–199.
- J. BYBEE (1985). *Morphology : a study of the relation between meaning and form*. Typological studies in language. John Benjamins.

- Kersti BÖRJARS, Nigel VINCENT, Carol CHAPMAN (1996). “Paradigms, periphrases and pronominal inflection : a feature-based account”. In : Geert BOOIJ, Jaap VAN MARLE (éds.), *Yearbook of Morphology 1996*, Dordrecht : Kluwer, pp. 155–180.
- Antoinette CAMILLERI, Martine VANHOVE (1994). “A phonetic and phonological description of the Maltese dialect of Mġarr (Malta)”. In : *Zeitschrift für arabische Linguistik 28*, pp. 87–110.
- Maris CAMILLERI (2011). “Island Morphology : Morphology’s Interactions In the Study of Stem Patterns”. In : *Linguistica 51*, pp. 65–84. Internal and External Boundaries of Morphology.
- (2012). “Stem patterns across Maltese verbal paradigms”. Présentation au 15<sup>ème</sup> International Morphology Meeting (IMM15). Du 9 au 12 janvier 2012. Vienne, Autriche.
- Maris CAMILLERI, Géraldine WALTHER (2012). “What small vowels and a large lexicon tell us about Maltese verbal inflection”. Présentation au 8<sup>ème</sup> Colloque des Décembrettes (Décembrettes 8). Du 6 au 7 décembre 2012. Bordeaux, France.
- Peter CARRUTHERS (2002). “The cognitive functions of language”. In : *Behavioral and Brain Sciences 25*, pp. 657–726.
- Shirley CHAPMAN, Desmond C. DERBYSHIRE (1991). “Paumari”. In : Desmond C. DERBYSHIRE, Geoffrey K. PULLUM (éds.), *Handbook of Amazonian Languages 3*, Berlin : Mouton de Gruyter, pp. 161–352.
- Noam CHOMSKY (1968). *Language and mind*. New York : Harcourt, Brace & World.
- (1970). “Remarks on Nominalization”. In : R. JACOBS, P. ROSENBAUM (éds.), *Readings in English Transformational Grammar*, Waltham, Massachusetts, États Unis d’Amérique : Ginn, pp. 184–221.
- (1981). *Lectures on Government and Binding*. Dordrecht, Pays-Bas : Foris.
- Marcus Tullius Cicero (dit) CICÉRON (52 AEC). “Pro Tito Annio Milone ad iudicem oratio (Pro Milone)”.
- Lionel CLÉMENT, Benoît SAGOT, Bernard LANG (2004). “Morphology Based Automatic Acquisition of Large-coverage Lexica”. In : *Actes de LREC’04*. Lisbonne, Portugal, pp. 1841–1844.
- Bernard COMRIE, Ray FABRI, Elizabeth HUME, Manwel MIFSUD, Thomas STOLZ, Martine VANHOVE (éds.) (2009). *Introducing Maltese Linguistics : Selected Papers from the 1st International Conference on Maltese Linguistics*. Amsterdam, Pays-Bas : John Benjamins. 18-20 October 2007, Bremen, Germany.



- Greville G. CORBETT (2003). “Agreement : the range of the phenomenon and the principles of the Surrey database of agreement”. In : *Transactions of the philological society* 101, p. 155–202.
- (2007a). “Canonical Typology, suppletion and possible words”. In : *Language* 83, pp. 8–42.
- (2007b). “Deponency, Syncretism, and What Lies Between”. In : Matthew BAERMAN, Greville G. CORBETT, Dunstan BROWN, Andrew HIPPISEY (éds.), *Deponency and Morphological Mismatches*. The British Academy, Oxford University Press, t. 145, pp. 21–43.
- (2010). “Morphomic Splits”. Présentation à l’atelier ”Perspectives on the Morphome” à l’Université de Coimbra.
- Greville G. CORBETT, Matthew BAERMAN (2006). “Prolegomena to a typology of morphological features”. In : *Morphology* 16, 2, pp. 231–246.
- Greville G. CORBETT, Norman FRASER (1993). “Network Morphology : a DATR account of Russian nominal inflection”. In : *Journal of Linguistics* 29, pp. 113–142.
- Jan Niecisław Ignacy BAUDOIN DE COURTENAY (1895). *Versuch einer Theorie phonetischer Alternationen : Ein Kapitel aus der Psychophonetik*. Strasbourg, Empire allemand.
- Blandine COURTOIS (1995). “Buts et méthodes de l’élaboration des dictionnaires électroniques du LADL”. In : *Théories et pratiques du lexique*, Université Paris 7, Cahiers du CIEL.
- Denis CREISSELS (2006). *Syntaxe générale — une introduction typologique 1 : catégories et constructions*. Paris, France : Hermès-Lavoisier.
- Berthold CRYSMANN (2002). *Constraint-Based Coanalysis : Portuguese Cliticisation and Morphology-Syntax Interaction in HPSG*. N° 15 in Saarbrücken Dissertations in Computational Linguistics and Language Technology. Saarbrücken : Deutsches Forschungszentrum für Künstliche Intelligenz and Universität des Saarlandes.
- Mary DALRYMPLE (2001). *Lexical Functional Grammar*, t. 34 de *Syntax and Semantics*. Academic Press.
- Anna-Maria DI SCIULLO, Edwin WILLIAMS (1987). *On the Definition of Word*. Cambridge, MA, États Unis d’Amérique : MIT Press.
- Lise DOBRIN (2001). “Arapesh”. In : Jane GARRY, Carl RUBINO (éds.), *Encyclopedia of the World’s Languages : Past and Present*, New York and Dublin : H. W. Wilson Press, pp. 33–38.
- Guillaume D’OKHAM (1323). *Summa totius logicae*, Éditions T.E.R, t. I, chap. 12. 1988 éd.

- Wolfgang U. DRESSLER (1990). “The cognitive perspective of ”naturalist” linguistic models”. In : *Cognitive Linguistics* 1, 1, pp. 75–98.
- (2000). “Naturalness”. In : G.E. BOOIJ, C. LEHMANN, J. MUGDAN (éds.), *Morphologie/Morphology : Ein Internationales Handbuch Zur Flexion Und Wortbildung/ an International Handbook on Inflection and Word-Formation*, Walter de Gruyter, Handbücher zur Sprach- und Kommunikationswissenschaft, pp. 288–296.
- Wolfgang U. DRESSLER, Lavinia MERLINI BARBARESI (1994). *Morphopragmatics*. de Gruyter.
- Matthew S. DRYER (1998). “Aspects of Word Order in the Languages of Europe”. In : Anna SIEWIERSKA (éd.), *Constituent Order in the Languages of Europe*, Berlin : Mouton de Gruyter, European Science Foundation Language Typology series, pp. 283–319.
- Matthew S. DRYER, Martin HASPELMATH (éds.) (2011). *The World Atlas of Language Structures Online*. Munich : Max Planck Digital Library. Consulté en été 2012, URL <http://wals.info/>.
- Michael DUNN (1999). *A Grammar of Chukchi*. Thèse de doctorat, Australian National University.
- Alfred ERNOUT, François THOMAS (1953). *Syntaxe Latine*. Paris : Klincksieck, 2 éd.
- Nicholas EVANS (1995). *A Grammar of Kayardild. With Historical-Comparative Notes on Tangkic*, t. 15 de *Mouton Grammar Library*. Berlin : Mouton de Gruyter.
- Nicholas EVANS, Stephen C. LEVINSON (2009). “The myth of language universals : Language diversity and its importance for cognitive science”. In : *Behavioral and Brain Sciences* 32, pp. 429–492.
- Roger P. EVANS, Gerald GAZDAR (1989). “Inference in DATR”. In : *Proceedings of the Fourth Conference of the European Chapter of the Association for Computational Linguistics*. pp. 66–71.
- (1996). “DATR : a language for lexical knowledge representation”. In : *Computational Linguistics* 22, pp. 167–216.
- Ray FABRI (2009). “Stem allomorphy in the Maltese verb”. In : *Ilsienna - Our Language* 1, pp. 1–20.
- Raphael FINKEL, Gregory T. STUMP (2002). “Generating Hebrew Verb Morphology by Default Inheritance Hierarchies”. In : *Proceedings of the ACL-02 Workshop on Computational Approaches to Semitic Languages*. Philadelphia, Pennsylvania, États Unis d’Amérique : Association for Computational Linguistics.
- Pierre FLOBERT (1967). “Déponent et passif en italique et en celtique”. In : *Annales de Bretagne* 74, 4, pp. 567–604.

- (1975). *Les Verbes déponents latins des origines à Charlemagne*. NS Recherches' 17. Paris, France : Publications de la Sorbonne, les belles lettres éd.
- Bernard FRADIN (2003). *Nouvelles approches en morphologie*. Paris, France : Presses Universitaires de France.
- Bernard FRADIN, Françoise KERLEROUX (2003). “Troubles with Lexemes”. In : Geert BOOIJ, Angela Ralli JANET DE CESARIS, Sergio Scalise (éds.), *Selected papers from the Third Mediterranean Morphology Meeting*. Barcelone, Espagne : IULA-Universitat Pompeu Fabra, Topics in Morphology, pp. 177–196.
- F. GAFFIOT, P. FLOBERT (2000). *Le Grand Gaffiot : dictionnaire latin-français*. Hachette-Livre.
- Gerald GAZDAR, Ewan KLEIN, Geoffrey PULLUM, Ivan SAG (1985). *Generalized Phrase Structure Grammar*. Basil Blackwell.
- Michel de GIVE (1992). *Grammaire latine*. De Boeck Secondaire.
- P.G.W. GLARE (1983). *Oxford Latin Dictionary*. Oxford Latin Dictionary. Oxford, Rauyaume Uni : Oxford University Pres.
- John GOLDSMITH (2001). “Unsupervised Learning of the Morphology of a Natural Language”. In : *Computational Linguistics* 27, 2, pp. 153–198.
- Henri GRAPPIN (1985). *Grammaire de la langue polonaise*. Paris : Institut d'études slaves.
- Joseph H. GREENBERG (1963). “Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements”. In : *Universals of Language*, MIT Press, pp. 73–113.
- (1966). “Synchronic and diachronic universals in phonology”. In : *Language* 42, 2, p. 508–17.
- Maurice GROSS (1975). *Méthodes en syntaxe*. Paris, France : Hermann.
- K. S. GURUBASAVE GOWDA (1975). *Ao Grammar*. Manasagangotri, Mysore : Central Institute of Indian Languages.
- Roy Stephen HAGMAN (1977). *Nama Hottentot Grammar*. Bloomington : Indiana University Press.
- Geoffrey HAIG (2010). *Alignment change in Iranian languages : a construction grammar approach*. De Gruyter.
- Morris HALLE (1973). “Prolegomena to a Theory of Word Formation”. In : *Linguistic Inquiry* 4, 1, pp. 3–16.

- Morris HALLE, Alex MARANTZ (1993). “Distributed morphology and the pieces of inflection”. In : Kenneth HALE, Samuel Jay KEYSER (éds.), *The view from building 20*, Cambridge, États Unis d’Amérique : MIT Press, pp. 111–176.
- Heather K. HARDY, Timothy MONTLER (1988). “Imperfective Gemination in Alabama”. In : *International Journal of American Linguistics* 54, pp. 399–475.
- Alice HARRIS (2002). *Endoclitics and the Origins of Udi Morphosyntax*. Oxford, Royaume-Uni : Oxford University Press.
- (2009). “Exuberant Exponence in Batsbi”. In : *Natural Language and Linguistic Theory* 27, 267, pp. 267–303.
- Andrew HIPPISEY (1997). *Declarative derivation*. Thèse de doctorat, University of Surrey.
- (2007). “Declarative Deponency : A Network Morphology Account of Morphological Mismatches”. In : Matthew BAERMAN, Greville G. CORBETT, Dunstan BROWN, Andrew HIPPISEY (éds.), *Deponency and Morphological Mismatches*. The British Academy, Oxford University Press, t. 145, pp. 145–173.
- Louis HJELMSLEV (1966). *Le langage : une introduction*. Paris : Éditions de Minuit. Traduction de M. Olsen.
- Robert HOBERMAN, Mark ARONOFF (2002). “The verbal morphology of Maltese : From Semitic to Romance”. In : Joseph SHIMRON (éd.), *Root-Based Morphology*, Amsterdam, Pays-Bas : John Benjamins.
- Charles F. HOCKETT (éd.) (1970). *A Leonard Bloomfield Anthology*. Bloomington, États Unis d’Amérique : Indiana University Press.
- Charles Francis HOCKETT (1954). “Two models of linguistic descriptions”. In : *Words* 10, pp. 210–234.
- Nancy IDE, Jean VÉRONIS (1994). “MULTTEXT : Multilingual Text Tools and Corpora”. In : *Proceedings of the 14th Conference on Computational Linguistics (COLING’94)*. Kyoto, Japon.
- Aleksandr V. ISAČENKO (1962). *Die russische Sprache der Gegenwart*, t. Teil I. Formenlehre. Halle (Saale), Allemagne : Niemeyer.
- Ray JACKENDOFF (1975). “Morphological and Semantic Regularities in the Lexicon”. In : *Language* 51, 3, pp. 639–671.
- Roman JAKOBSON (1965). *Selected Writings*, t. II. La Haye, Pays-Bas : Mouton.
- Patrick JUOLA (2008). “Assessing Linguistic Complexity”. In : Matti MIESTAMO, Kaius SINNEMÄKI, Fred KARLSSON (éds.), *Language Complexity : Typology, Contact, Change*, Amsterdam, Pays-Bas : John Benjamins Press.

- Lauri KARTTUNEN (2003). "Computing with Realizational Morphology". In : Alexander GELBUKH (éd.), *Computational Linguistics and Intelligent Text Processing*, Heidelberg, Allemagne : Springer Verlag, t. 2588 de *Lecture Notes in Computer Science*, pp. 205–216.
- Александр Е. Кибрик — Alexander E KIBRIK (1977). *Опыт структурного описания Арчинского языка Том 2 : таксономическая грамматика (Taksonomičeskaja grammatika. vol. 2 of Opyt strukturnogo opisanija arčinskogo jazyka)*. Moscou, URSS : Издательство Московского Университета (Izdatel'stvo moskovskogo universiteta).
- Marianne KILANI-SCHOCH, Wolfgang U. DRESSLER (2005). *Morphologie naturelle et flexion du verbe français*. Tübingen, Allemagne : Gunter Narr Verlag Tübingen.
- Geoffrey D. KIMBALL (1988). "Koasati Reduplication". In : William SHIPLAY (éd.), *In Honor of Mary Haas : from the Haas Festival Conference on Native American Linguistics*, Berlin, Allemagne : Mouton de Gruyter, pp. 431–442.
- Paul KIPARSKY (1973). "'Elsewhere' in Phonology". In : Stephen ANDERSON, Paul KIPARSKY (éds.), *A Festschrift for Morris Halle*, New York, NY, États Unis d'Amérique : Holt, Rinehart and Winston, pp. 93–106.
- (1982). "Word-formation and the Lexicon". In : F. INGEMANN (éd.), *Proceedings of the Mid-America Linguistics Conference*. Lawrence, Kansas, États Unis d'Amérique.
- (2005). "Blocking and periphrasis in inflectional paradigms". In : *Yearbook of Morphology 2004*. Dordrecht : Springer, pp. 113–135.
- Jean-Pierre KOENIG (1999). *Lexical Relations*. Stanford Monographs in Linguistics. Center for the Study of Language and Inf.
- Kimmo KOSKENNIEMI (1984). "A general computational model for word-form recognition and production". In : *Proceedings of the 22nd annual meeting of the Association for Computational Linguistics*. pp. 178–181.
- A. L. КРОЕВЕР, George William GRACE (1960). *The Sparkman Grammar of Luiseño*, t. 16 de *University of California Publications in Linguistics*. Berkeley : University of California Press.
- Елена С. Кубрякова — Elena S. KUBRЯKOVA (1972). "О соотношении парадигматических и словообразовательных рядов в германских языках (O sootnošenii paradigmatičeskix i slovoobrazovatelnyx rjadov v germanskix jazykax)". In : *Историко-типологические исследования морфологического строя германских языков*, Moscou, URSS : Наука, pp. 172–188.
- Aert H. KUIPERS (1967). *The Squamish Language, Grammar, Texts, Dictionary*, t. 73 de *Janua Linguarum, Series Practica*. The Hague : Mouton.
- G. LAZARD, C. MOYSE-FAURIE (2005). *Linguistique typologique. Sens et Structures*. Presses Universitaires du Septentrion.

- Gilbert LAZARD (2003). Présentation au International Symposium on Linguistic Diversity and Language Theories.
- Gilbert LAZARD, Yann RICHARD, Rokhsareh HECHMATI, Pollet SAMVELIAN (2006). *Grammaire du persan contemporain*. Téhéran, Iran : Institut Français de Recherche en Iran & Farhang Moaser Edition.
- Charlton T. LEWIS, Charles SHORT (1879). *A Latin Dictionary*. Oxford, Royaume Uni : Clarendon Press. Disponible en ligne sur le site du projet Perseus.
- Rochelle LIEBER (1981). *On the Organization of the Lexicon*. Thèse de doctorat, University of New Hampshire.
- (1992). *Deconstructing Morphology : Word Formation in Syntactic Theory*. Chicago, États Unis d'Amérique : University of Chicago Press.
- Jean LOWENSTAMM (1996). "CV as the Only Syllable Type". In : Jacques DURAND, Bernard LAKS (éds.), *Current Trends in Phonology Models and Methods*, European Studies Research Institute, University of Salford, pp. 419–442.
- Martin MAIDEN (1992). "Irregularity as a determinant of morphological change". In : *Journal of Linguistics* 28, p. 285–312.
- (2012). "A possible constraint on the domain of morphemes : The morphology of Romanian feminine nouns." Présentation au 15ème International Morphology Meeting (IMM15). Du 9 au 12 janvier 2012. Vienne, Autriche.
- Robert MALOUF (2005). "Disjunctive rule ordering in finite state morphology". Communication au 41st Meeting of the Chicago Linguistics Society.
- Oskar MANN (1906). *Die Mundarten der Mukri-Kurden, Teil I: Grammatikalische Skizze, Texte in phonetischer und persische Umschrift*. Berlin, Allemagne : Georg Reimer.
- Alec MARANTZ (1997). "No escape from syntax : Don't try morphological analysis in the privacy of your own Lexicon". In : Alexis DIMITRIADIS, Laura SIEGEL, Clarissa SUREK-CLARK, Alexander WILLIAMS (éds.), *Proceedings of the 21st Annual Penn Linguistics Colloquium*. Penn Working Papers in Linguistics, pp. 201–225.
- Peter H. MATTHEWS (1972). *Inflectional Morphology : a Theoretical Study Based on Aspect of Latin Verb Conjugation*. Cambridge, Royaume Uni : Cambridge University Press.
- (1974). *Morphology*. Cambridge, Royaume Uni : Cambridge University Press.
- Willi MAYERHALER (1981). *Morphologische Natürlichkeit*. Athenaiion. Traduction anglaise (1988) : Morphological Naturalness. Ann Arbor : Karoma Publ.
- John J MCCARTHY (1981). "A prosodic theory of nonconcatenative morphology". In : *Linguistic Inquiry* 12, pp. 373–418.

- John J. McCARTHY (1982). “Prosodic structure and expletive infixation.” In : *Language* .
- Ernest N. McCARUS (1958). *A Kurdish Grammar : descriptive analysis of the Kurdish of Sulaimaniya, Iraq*. Thèse de doctorat, American Council of Learned Societies, New-York, États Unis d’Amérique.
- John McWHORTER (2001). “The world’s simplest grammars are creole grammars”. In : *Linguistic Typology* 5, pp. 125–166.
- Neil MERCER (2000). *Words and Minds. How We Use Language to Think Together*. Londres, Royaume-Uni : Routledge.
- Gabriele MICELLI, Alfonso CARAMAZZA (1989). “Dissociation of inflectional and derivational morphology”. In : *Brain and Language* 35, 1, pp. 24–65.
- Manwel MIFSUD (1995). *Loan verbs in Maltese*. Brill.
- Miguel Ángel MOLINERO, Benoît SAGOT, Lionel NICOLAS (2009). “A morphological and syntactic wide-coverage lexicon for Spanish : The Leffe”. In : *Proceedings of the 7th conference on Recent Advances in Natural Language Processing (RANLP 2009)*. Borovets, Bulgarie.
- Richard MONTAGUE (1970). “English as a Formal Language”. In : *Linguaggi nella società e nella tecnica* pp. 189–224.
- Stanley NEWMAN (1944). *The Yokuts Language of California*. New York : The Viking Fund Publications in Anthropology.
- Rachel NORDLINGER (1998). *Constructive Case*. Thèse de doctorat, Stanford University. Revised version of PhD thesis.
- Rachel NORDLINGER, Louisa SADLER (2004). “Relating Morphology to Syntax”. In : Louisa SADLER, Andrew SPENCER (éds.), *Projecting Morphology*, Stanford, États Unis d’Amérique : CSLI Publications, pp. 159–185.
- Enrique PALANCAR (). “An alternative view on the inflectional complexity of Chinantecan verb classes”. In : *TBA* à paraître.
- Carole PARADIS (1988). “On Constraints and Repair Strategies”. In : *The Linguistic Review* 6, 1, pp. 71–97.
- Carole PARADIS, Fatimazohra El FENNE (1995). “French verbal inflection revisited : Constraints, repairs and floating consonants”. In : *Lingua* 95, pp. 169–204.
- PĀṆINI (env. 500 AEC). “Aṣṭādhyāyī : Grammaire du sanskrit”.
- Charles S. PEIRCE (1965). *Collected Papers*. Cambridge, MA, États Unis d’Amérique : Harvard University Press.

- David PERLMUTTER (1988). “The Split Morphology Hypothesis : Evidence from Yiddish”. In : Michael HAMMOND, Michael NOONAN (éds.), *Theoretical Morphology*, San Diego, États Unis d’Amérique : Academic Press, Inc., pp. 79–99.
- David M. PERLMUTTER (1980). “Relational grammar”. In : E. A. MORAVCSIK, J. R. WIRTH (éds.), *Syntax and semantics : Current approaches to syntax*, New York : Academic Press.
- Steven T. PIANTADOSI, Edward GIBSON (2011). “Quantitative standards for absolute linguistic universals”. In : *under review* .
- Vito PIRRELLI, Marco BATTISTA (2000). “The Paradigmatic Dimension of Stem Allomorphy in Italian Verb Inflection”. In : *Italian Journal of Linguistics* pp. 307–380.
- Marc PLÉNAT (1987). “Morphologie du passé simple et du passé composé des verbes de l’og autre” conjugaison”. In : *ITL Review of Applied Linguistics* 77, 78, pp. 93–150.
- Alan PRINCE, Paul SMOLENSKY (1993). “Optimality Theory : Constraint Interaction in Generative Grammar”. Rap. tech., Rutgers University Center for Cognitive Science and Computer Science Department, University of Colorado at Boulder. Blackwell Publishers (2004).
- Richard E. PRIOR, Joseph WOHLBERG (1995). *501 Latin verbs — fully conjugated in all the tenses*. Barron’s Educational Series, Inc. Barron’s.
- J. RISSANEN (1984). “Universal coding, information, prediction, and estimation”. In : *IEEE Transactions on Information Theory* 30, 4, pp. 629–636.
- Brian ROARK, Richard SPROAT (2007). *Computational Approaches to Morphology and Syntax*. Oxford Surveys in Syntax and Morphology. Oxford University Press, États Unis d’Amérique.
- Robert H. ROBINS (1959). “In defense of WP”. In : *Transactions of the Philological Society* 1959 pp. 116–144.
- Michel ROCHÉ (2010). “Base thème, radical”. In : *Recherches linguistiques de Vincennes* 39.
- Carl RUBINO (2000). *Ilocano Dictionary and Grammar*. Honolulu : University of Hawaii Press.
- (2001). “Pangasinan”. In : Jane GARRY, Carl RUBINO (éds.), *Encyclopedia of the World’s Languages : Past and Present*, New York / Dublin : H.W. Wilson Press, pp. 539–542.
- (2011). “Reduplication”. In : Matthew S. DRYER, Martin HASPELMATH (éds.), *The World Atlas of Language Structures Online*, Munich : Max Planck Digital Library. Chapitre 27, accessible en ligne. Consulté en été 2012, URL <http://wals.info/chapter/27>, .



- Louisa SADLER, Andrew SPENCER (1998). “Argument structure and morphology”. In : Andrew SPENCER, Arnold ZWICKY (éds.), *The Handbook of Morphology*, Oxford, Royaume Uni : Basil Blackwell, pp. 206–36.
- (2001). “Syntax as an exponent of morphosyntactic features”. In : Geert BOOIJ, Jaap VAN MARLE (éds.), *Yearbook of Morphology 2000*, Dordrecht, Pays-Bas : Kluwer Academic Publishers, pp. 71–96.
- Ivan SAG, Thomas WASOW, Emily BENDER (2003). *Syntactic Theory : a Formal Introduction*. Palo Alto, Californie : CSLI.
- Benoît SAGOT (2005). “Automatic acquisition of a Slovak lexicon from a raw corpus”. In : *Lecture Notes in Artificial Intelligence 3658 ((c) Springer-Verlag), Proceedings of TSD’05*. Karlovy Vary, République Tchèque, p. 156–163.
- (2006). *Analyse automatique du français : lexiques, formalismes, analyseurs*. Thèse de doctorat, Université Paris 7.
- (2007). “Building a morphosyntactic lexicon and a pre-syntactic processing chain for Polish”. In : *Proceedings of the 3rd Language & Technology Conference (LTC’05)*. Poznań, Pologne, p. 423–427.
- (2010). “The Lefff, a freely available, accurate and large-coverage lexicon for French”. In : *Proceedings of the 7th Language Resource and Evaluation Conference*. La Valette, Malte.
- Benoît SAGOT, Lionel CLÉMENT, ÉRIC DE LA CLERGERIE, Pierre BOULLIER (2006). “The Lefff 2 syntactic lexicon for French : architecture, acquisition, use”. In : *Proceedings of the 5th Language Resource and Evaluation Conference*. Lisbonne, Portugal.
- Benoît SAGOT, Géraldine WALTHER (2010). “Développement de ressources pour le persan : lexique morphologique et chaîne de traitements de surface”. In : *Actes de TALN 2010*. Montréal, Canada.
- (2011). “Non-canonical inflection : data, formalisation and complexity measures.” In : Cerstin MAHLOW, Michael PIOTROWSKI (éds.), *Systems and Frameworks in Computational Morphology*. Zurich, Suisse : Springer, t. 100, pp. 23–45.
- (2013). “Implementing a formal model of inflectional morphology”. In : Cerstin MAHLOW, Michael PIOTROWSKI (éds.), *Systems and Frameworks in Computational Morphology*. Berlin, Heidelberg, Allemagne : Springer, t. 100.
- Benoît SAGOT, Géraldine WALTHER, Pegah FAGHIRI, Pollet SAMVELIAN (2011). “Développement de ressources pour le persan : PerLex 2 et MElt<sub>fa</sub>”. In : *Actes de TALN 2011*. Montpellier, France.
- Pollet SAMVELIAN (2007). “What Sorani Kurdish Absolute Prepositions Tell us About Cliticization”. In : Frederic HOYT, Nikki SEIFERT, Alexandra TEODORESCU (éds.),

- Texas Linguistics Society IX : The Morphosyntaxe of Understudied Languages*. Stanford, États Unis d'Amérique : CSLI online publications.
- Lucien SAUSY (1977). *Grammaire Latine complète*. Fernand Lanore.
- Anna SIEWIERSKA (2012). "Passive Constructions". In : Matthew S. DRYER, Martin HASPELMATH (éds.), *The World Atlas of Language Structures Online*, Munich, Allemagne : Max Planck Digital Library, chap. 1. Chapitre 107, accessible en ligne. Consulté en été 2012, URL <http://wals.info/chapter/1>.
- Michael SPAGNOL (2011). *A Tale of Two Morphologies. Verb Structure and argument alternations in Maltese*. Thèse de doctorat, University of Konstanz, Constance, Allemagne.
- David STAMPE (1973). "On Chapter Nine". In : Michael KENSTOWICZ, Charles KISSEBERTH (éds.), *Issues in Phonological Theory*, La Haye, Pays-Bas : Mouton, pp. 44–52.
- Susan STEELE (1995). "Towards a Theory of Morphological Information". In : *Language* 71, 2, pp. 260–309.
- Gregory T. STUMP (1991). "A Paradigm-theory of Morphosemantic Mismatches". In : *Language* 67, p. 675–725.
- (2001). *Inflectional Morphology. A Theory of Paradigm Structure*. Cambridge, Royaume Uni : Cambridge University Press.
- (2002). "Morphological and syntactic paradigms : Arguments for a theory of paradigm linkage". In : Geert BOOIJ, Jaap VAN MARLE (éds.), *Yearbook of Morphology 2001*, Kluwer, pp. 147–180.
- (2006). "Heteroclisys and Paradigm Linkage". In : *Language* 82, pp. 279–322.
- Pierre SWIGGERS, Karel VAN DEN EYNDE (1987). "La morphologie du verbe français". In : *ITL Review of Applied Linguistics* 78, pp. 151–251.
- W.M. THACKSTON (2006). "Sorani Kurdish : A Reference Grammar with Selected Readings". Publié en ligne.
- Anna M. THORNTON (2010). "Towards a Typology of overabundance". Communication au colloque des Décembrettes 7.
- (2011). "Overabundance (Multiple Forms Realizing the Same Cell) : A Non-canonical Phenomenon in Italian Verb Morphology". In : Maria Goldbach MARTIN MAIDEN, John Charles Smith, Marc-Olivier HINZELIN (éds.), *Morphological Autonomy: Perspectives From Romance Inflectional Morphology*, Oxford University Press.
- (2012). "Reduction and maintenance of overabundance. A case study on Italian verb paradigms". In : *Word Structure* 5, pp. 183–207.

- Mauro Tosco (2001). *The Dhaasanac language*. Cologne, Allemagne : Rüdiger Klöppe.
- Delphine TRIBOUT (2010). *Les conversions de nom à verbe et de verbe à nom en français*. Sciences du langage, Université Paris Diderot (Paris 7), Paris, France.
- A USSISHKIN, A TWIST (2009). “Auditory and visual lexical decision in Maltese”. In : Bernard COMRIE, Ray FABRI, Elizabeth HUME, Manwel MIFSUD, Thomas STOLZ, Martine VANHOVE (éds.), *Introducing Maltese Linguistics : Proceedings of the 1st International Conference on Maltese Linguistics*. John Benjamins, Studies in Language Companion Series, pp. 233– 249.
- Dale F. WALKER (1976). *Grammar of the Lampung Language*. Jakarta : NUSA.
- Géraldine WALTHER (2011a). “A Derivational Account for Sorani Kurdish Passives.” Communication au colloque 4th International Conference on Iranian Linguistics (ICIL4).
- (2011b). “Measuring Morphological Canonicity”. In : *Linguistica 51*, pp. 157–180. Internal and External Boundaries of Morphology.
- (2012a). “Fitting into morphological structure : accounting for Sorani Kurdish endoclitics”. In : *Proceedings of the 8th Mediterranean Morphology Meeting*. Cagliari, Italie.
- (2012b). “Formally Encoding Morphomicity”. Affiche au 15ème International Morphology Meeting (IMM15). Du 9 au 12 janvier 2012. Vienne, Autriche.
- Géraldine WALTHER, Benoît SAGOT (2010). “Developing a Large-Scale Lexicon for a Less-Resourced Language : General Methodology and Preliminary Experiments on Sorani Kurdish”. In : *Proceedings of the 7th SaLTMiL Workshop on Creation and use of basic lexical resources for less-resourced languages (LREC 2010 Workshop)*. La Vallette, Malte.
- (2011). “Modélisation et implémentation de phénomènes non-canoniques”. In : *Traitement Automatique des Langues 52*, 2.
- Géraldine WALTHER, Benoît SAGOT, Karën FORT (2010). “Fast Development of Basic NLP Tools : Towards a Lexicon and a POS Tagger for Kurmanji Kurdish”. In : *Actes du 29<sup>e</sup> Colloque sur le Lexique et la Grammaire*. Belgrade, Serbie.
- Edwin WILLIAMS (1981). “On the notions ’lexically related’ and ’head of word’”. In : *Linguistic Inquiry*.
- Wolfgang U. WURZEL (1984). *Flexionsmorphologie und Natürlichkeit*. Berlin, Allemagne : Akademie-Verlag. Traduction anglaise : Wurzel (1989).
- (1989). *Inflectional Morphology and Naturalness*. Studies in Natural Language and Linguistic Theory. Kluwer Academic Publishers.

- Aris XANTHOS (2008). *Apprentissage automatique de la morphologie — Le cas des structures racine-schème*, t. 48 de *Sciences pour la Communication*. Peter Lang.
- Zheng XU, Mark ARONOFF, Frank ANSHEN (2007). *Deponency and morphological mismatches*, Oxford, Rauyaume Uni : Oxford University Pres, t. 145 de *Proceedings of the British Academy*, chap. Deponency in Latin, pp. 127–143.
- Alfonz ZAUNER (1973). *Praktická príručka slovenského pravopisu*. Martin, Slovaquie : Vydavateľstvo Osveta.
- Arnold M. ZWICKY (1985). “How to describe inflection”. pp. 372–386.
- Arnold M. ZWICKY, Geoffrey K. PULLUM (1988). “The syntax-phonology interface”. In : Frederick J. NEWMAYER (éd.), *Linguistic Theory : Foundations*, Cambridge, États Unis d’Amérique : Cambridge University Press, Linguistics : The Cambridge Survey, chap. 10, pp. 255–280.

