



THÈSE

**En vue de l'obtention du
DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE**

Délivré par l'Université Toulouse 2 - Jean Jaurès

Présentée et soutenue par

Daniele SANACORE

Le 7 décembre 2023

**Une histoire de famille: description morphosémantique des
lexèmes construits et des relations dérivationnelles**

Ecole doctorale : **CLESCO - Comportement, Langage, Education, Socialisation,
Cognition**

Spécialité : **Sciences du langage**

Unité de recherche :

CLLE - Unité Cognition, Langues, Langage, Ergonomie

Thèse dirigée par

Nabil HATHOUT et Fiammetta NAMER

Jury

Mme Georgette DAL, Rapporteur

M. Richard HUYGHE, Rapporteur

Mme Stéphanie LIGNON, Examinatrice

M. Nabil HATHOUT, Directeur de thèse

Mme Fiammetta NAMER, Co-directrice de thèse

M. Fabio MONTERMINI, Président

Remerciements

Le mot “histoire”, comme le titre de la thèse le suggère, apparaîtra tout au long de ce manuscrit. Je me permets de l’utiliser une fois de plus pour exprimer ma gratitude envers celles et ceux que j’ai eu le privilège de croiser au cours de cette longue histoire qui a été ma thèse.

Tout d’abord, un énorme remerciement à mes directeurs de thèse Nabil Hathout et Fiammetta Namer. Merci de m’avoir offert l’opportunité d’entreprendre cette thèse, pour votre disponibilité et patience infinie, pour toutes vos relectures, tous vos conseils de ces années et pour m’avoir accompagné tout au long de ce travail.

Je souhaite également remercier Georgette Dal et Richard Huyghe d’avoir accepté d’être les rapporteurs de cette thèse, ainsi que Stéphanie Lignon et Fabio Montermini pour avoir accepté de faire partie du jury.

Je tiens à remercier les membres permanents et le personnel administratif du laboratoire CLLE pour leur accueil et pour m’avoir fourni tous les moyens nécessaires pour réaliser mon travail dans les meilleures conditions possibles.

Une mention spéciale pour les nombreux collègues doctorants et postdocs que j’ai pu rencontrer dans mon parcours. On peut parfois éprouver une certaine solitude au cours d’une thèse, mais grâce à votre soutien et votre amitié, je ne me suis jamais retrouvé dans cette situation. Vous êtes vraiment nombreux et même si je devrais consacrer plusieurs pages pour remercier chacun d’entre vous individuellement comme vous le méritez, je vais devoir me limiter à vous citer dans ce paragraphe et je m’en excuse. Merci donc à Efisio Manunta, Marc-Philippe Brunet, Océane Abrard, Silvia Federzoni, Chiara Minoccheri, Maxime Delmas, Valentin Goulette, Robin Cazes, Quentin Marre, Florian Alba, Dimitri Paisios, Claire Ibarboure, Mariame Maarouf, Malvina Brunet, Filip Miletic, Lison Fabre, Bénédicte Pierrejean, Marine Wauquier, Yizhe

Wang, Jasmine Abularach Mendoza, Killyam Forge, Clamença Poujade, Cécile Ferlin, Natalia Bobkova, Louise Garrigou, Victoria O'Callaghan, Yuting He Zhang, Brivael Hémon, Kevin De Checchi, Elodie Clayette, Alba Langlade, Carla Ganino, Lockman Aalioui, Emma Dagaud, Christophe Combe et tous les autres que je n'ai pas cités.

Un grand merci aux membres du département de Sciences du Langage pour m'avoir donné la possibilité de consacrer une partie de mon temps à l'enseignement et à tous les étudiants que j'ai rencontrés dans ces années.

Un grand merci à tous mes amis italiens de longue date qui m'ont toujours soutenu malgré la distance.

Pour conclure, les familles dérivationnelles sont sympas, mais je continue à préférer la mienne. Merci donc à ma famille pour leur amour inconditionnel et leur soutien constant.

Résumé

Cette thèse se consacre à l'étude du lexique construit français, à savoir la partie du lexique structurée par des relations morphologiques (ex. LAYER-LAVEUR), dans une perspective synchronique. En adoptant une approche paradigmatique des relations dérivationnelles (Bauer, 1997; Hathout et Namer, 2019; Ruz *et al.*, 2022), nous proposons une méthode générale et théorique pour organiser des familles de mots (i.e. familles dérivationnelles) en paradigmes.

L'approche que nous proposons est guidée par la sémantique. Nous considérons que les relations sémantiques entre les lexèmes contenus dans les familles sont premières pour la structuration des paradigmes en dérivation (Štekauer, 2014; Bonami et Strnadová, 2019). Nous délimitons les familles dérivationnelles sur la base des réseaux conceptuels dans lesquels nous pouvons les inscrire. Les familles structurées par le même réseau conceptuel peuvent être alignées dans des paradigmes.

Nous appelons les réseaux conceptuels dont nous nous servons *meaning bundles* (MB). Les MBs sont inspirés des scénarios cognitifs proposés par la Sémantique des Cadres de Fillmore (Fillmore *et al.*, 1976; Baker *et al.*, 1998). Les MBs se situent sur le plan conceptuel et contiennent des participants interprétables. Nous proposons d'obtenir les MBs à partir d'histoires qui décrivent les relations existant entre les lexèmes des familles.

Dans cette thèse, nous mettons en œuvre notre démarche sur des familles de nature sémantique variée, à savoir des familles basées sur des noms d'artefact (ex. POT, POTIER, POTERIE), des familles basées sur des noms de fruit (ex. BANANE, BANANIER, BANANERAIE) et des familles basées sur des noms d'animal (ex. LOUP, LOUVE, LOUVETEAU).

Mots clés : morphologie dérivationnelle, paradigmes, lexique construit, sémantique des cadres

Abstract

This thesis is devoted to the study of the French constructed lexicon, i.e. that part of the lexicon structured by morphological relations (e.g. LAYER-LAVEUR), from a synchronic perspective. By adopting a paradigmatic approach to derivational relations (Bauer, 1997; Hathout et Namer, 2019; Ruz *et al.*, 2022), we propose a general and theoretical methodology to organise word families (i.e. derivational families) into paradigms.

The approach that we propose is guided by semantics. We consider that semantic relations between lexemes contained in families are primary for the structure of derivational paradigms (Štekauer, 2014; Bonami et Strnadová, 2019). We delimit derivational families on the basis of the conceptual networks in which we can inscribe them. Families structured by the same conceptual network can be aligned in paradigms.

The conceptual networks that we use are called meaning bundles (MB). The MBs are inspired by the cognitive scenarios proposed by Fillmore’s Frame Semantics (Fillmore *et al.*, 1976; Petrucci, 1996). MBs are conceptually based and contain unpredictable participants. We propose to obtain MBs from stories that describe the relations between lexemes in derivational families. We use stories as heuristic tools to make relations between lexemes in families explicit.

In this thesis, we implement our approach on families of diverse semantic nature : families based on artefact nouns (e.g. POT, POTIER, POTERIE), families based on nouns denoting fruits (e.g. BANANE, BANANIER, BANANERAIE) and families based on nouns denoting animals (e.g. LOUP, LOUVE, LOUVETEAU).

Keywords : derivational morphology, paradigms, constructed lexicon, frame semantics

Table des matières

Résumé	v
Abstract	vii
Introduction	1
1 Système conceptuel et organisation du lexique	11
1.1 Introduction	12
1.2 Relations dérivationnelles et lexique construit	13
1.3 Familles et réseaux	15
1.3.1 Délimitation des familles dérivationnelles	16
1.3.2 Nature et fonctionnement des familles	17
1.4 Séries dérivationnelles	20
1.5 Bilan intermédiaire	23
1.6 Interface entre lexique et système conceptuel	23
1.6.1 Types de relations conceptuelles	25
1.6.2 Relations lexicales	26
1.7 Compositionnalité du sens lexical	28
1.8 Représentation des réseaux conceptuels	31
1.9 Bilan	32
2 Une approche sémantique aux paradigmes dérivationnels	35
2.1 Introduction	36

2.2	La notion de paradigme	38
2.2.1	Les paradigmes et leur usage en flexion	38
2.2.2	Adapter les paradigmes à la dérivation	40
2.2.3	À quoi ressemblent les paradigmes en dérivation?	41
2.2.4	Apport des paradigmes à l'analyse dérivationnelle	47
2.3	Exemples de modèles paradigmatiques	50
2.3.1	Les systèmes paradigmatiques de Bonami et Strnadová (2019)	50
2.3.2	Le modèle paradigmatique ParaDis	52
2.4	Variété sémantique des paradigmes en dérivation	55
2.4.1	Paradigme basé sur des noms de fruits	55
2.4.2	Paradigme basé sur les ethniques	57
2.4.3	Réseau action	59
2.4.4	Réseau activité	62
2.5	Bilan	64
3	Rôles thématiques et Sémantique des Cadres	67
3.1	Introduction	68
3.2	Les rôles sémantiques et leur usage en morphologie	69
3.3	Caractérisation sémantique des prédicats	74
3.4	Sémantique des Cadres et <i>FrameNet</i>	76
3.4.1	Modèle <i>FrameNet</i>	79
3.4.2	Rôles noyau	81
3.4.3	Caractérisation ontologique des rôles	82
3.4.4	Déclencheurs	83
3.4.5	Relations inter-cadre	85
3.5	Adaptabilité des cadres au découpage des familles	86
3.6	Bilan	87

4	Démarche <i>story-based</i>	89
4.1	Introduction à notre démarche	90
4.2	Données de départ et structures à produire	92
4.3	Organisation du lexique et structuration des paradigmes	95
4.4	Décrire la sémantique des relations morphologiques	95
4.5	Pertinence des modèles génératifs de langage	98
4.6	Formaliser le contenu des histoires	99
4.7	Variété du contenu sémantique des histoires	104
4.8	Aligner les relations contenues dans les histoires	109
4.9	Subset Meaning bundles	112
4.10	Traitement des autres sous-ensembles de la famille	118
4.11	Meaning bundles au niveau d’une famille	124
4.12	FMBs obtenus à partir d’autres familles et LMBs	127
4.13	Projection des LMBs sur le lexique	137
4.14	Bilan	142
5	Propriétés sémantiques des lexèmes et des paradigmes	145
5.1	Introduction	146
5.2	Relations sociales et comportement animal	147
5.2.1	Comportement social des animaux	149
5.2.2	Abri et groupe social	155
5.2.3	Proies et prédateurs	159
5.3	Animaux et activités humaines	163
5.3.1	Pêche	163
5.3.2	Chasse et animaux “nuisibles”	169
5.3.3	Élevage	176
5.3.4	Animaux en tant qu’objet d’étude	182
5.4	Bilan intermédiaire	185

5.5	Fruits	187
5.5.1	Production fruiticole	187
5.5.2	Transport spécialisé	191
5.5.3	Cuisine et recettes	195
5.6	Exemples de relation entre LMBs	199
5.7	Bilan	202
6	Relations périphériques	205
6.1	Introduction	205
6.2	Couples de noms d'humain masculin - féminin	207
6.3	Adjectifs relationnels	210
6.4	Adverbes en <i>-ment</i>	214
6.5	Noms d'action	214
6.6	Réitération d'un procès	215
6.7	Bilan	217
	Conclusions et perspectives	221
	Index	227
	bibliographie	243

Liste des tableaux

1	Exemple de paradigme en latin.	1
2	Exemple de paradigme dérivationnel	4
3	Paradigmes regroupant un extrait des familles de lexèmes qui interviennent dans une activité de fabrication	6
4	Paradigme dérivationnel regroupant des familles de lexèmes qui interviennent dans une activité de déplacement dans un récipient.	6
1.1	Exemples des différents types de familles et de séries proposés par Hathout (2009)	22
2.1	Paradigme abstrait d'un adjectif en français	38
2.2	Paradigme flexionnel de l'adjectif NATIONAL en français	39
2.3	Exemple de paradigme dérivationnel en français	44
2.4	Paradigme abstrait qui décrit les régularités dans le paradigme en Table 2.3 . . .	44
2.5	Exemple de paradigme dérivationnel structuré par des relations sémantiques . .	45
2.6	Exemple d'alignement de familles dérivationnelles "incomplètes"	46
2.7	Paradigme regroupant des familles contenant des noms dérivés en <i>-isme</i> and <i>-iste</i> mutuellement motivés (Hathout et Namer, 2019, p.7)	48
2.8	Paradigme fruit-plante-terrain	56
2.9	Exemple de paradigme dérivationnel basé sur les ethniques.	58
2.10	Paradigme dérivationnel structuré par le réseau action.	60
2.11	Des possibles sous-spécifications du réseau action	61
2.12	Exemple de paradigme dérivationnel structuré par le réseau activité proposé par Roché (2017b).	63

3.1	Définitions générales et exemples d'attribution des rôles sémantiques plus utilisés.	71
3.2	Liste de rôles sémantiques utilisés par Salvadori et Huyghe (2023) pour le typage des noms déverbaux.	73
4.1	Paradigmes regroupant un extrait des familles de lexèmes qui interviennent dans une activité de fabrication et qui ont comme élément initial une entité jouant le rôle d'OBJET FABRIQUÉ dans un LMB de fabrication.	140
4.2	Paradigmes regroupant un extrait des familles de lexèmes qui interviennent dans une activité de fabrication	141
4.3	Paradigme regroupant un extrait des familles de lexèmes qui interviennent dans une activité de fabrication et qui ont comme élément initial un nom dénotant un INSTRUMENT utilisé dans la fabrication.	141
4.4	Paradigme dérivationnel regroupant des familles de lexèmes qui interviennent dans une activité de déplacement dans un récipient.	142
5.1	Paradigme dérivationnel regroupant des familles contenant le mâle, la femelle et le petit d'un animal sauvage social	155
5.2	Paradigme structuré par les relations sémantiques existant entre un insecte, son abri et sa colonie.	159
5.3	Paradigme structuré par des relations de prédation entre espèces animales sauvages	162
5.4	Paradigme dérivationnel contenant des familles basées sur un nom d'ANIMAL PÊCHÉ et qui interviennent dans un scénario de pêche.	167
5.5	Paradigme dérivationnel regroupant des familles basées sur un INSTRUMENT DE PÊCHE.	169
5.6	Paradigmes dérivationnels structurés par une activité de chasse.	173
5.7	Paradigme dérivationnel qui regroupe des lexèmes inscrits dans une activité d'élimination d'animaux nuisibles.	176
5.8	Paradigme contenant des familles paradigmatiques qui rentrent dans un scénario d'élevage d'animaux en masse	180

5.9	Paradigme dérivationnel regroupant des familles paradigmatiques qui interviennent dans une activité d'élevage en ferme.	182
5.10	Paradigme dérivationnel regroupant des lexèmes intervenant dans une activité d'étude d'un animal.	185
5.11	Paradigme dérivationnel structuré par une activité de production fruiticole. . . .	190
5.12	Paradigme structuré par une activité de transport spécialisé d'une marchandise .	195
5.13	Paradigmes dérivationnel décrivant une préparation culinaire	198
5.14	Paradigme dérivationnel de familles basées sur un nom de préparation culinaire.	199
6.1	Organisation des noms d'humain au masculin et au féminin dans les paradigmes dérivationnels	210

Table des figures

1.1	Représentation sous forme d'arborescence d'une famille dérivationnelle partielle (Stump, 2019, p.3).	17
1.2	Représentation de la famille partielle de ENSEIGNER sous forme de graphe.	19
1.3	Concepts associés au concept CHAT sur le plan taxinomique et sur le plan engynomique.	26
1.4	Exemple de représentation de connaissance sous forme de réseau sémantique	32
2.1	Exemple d'alignement de trois paradigmes flexionnels concrets sous le même paradigme abstrait	40
2.2	Représentation sous-forme de réseau de la famille de HERBE	43
2.3	Patron permettant la description du sens de l'adjectif JOURNALISTIQUE	49
2.4	Alignement de trois familles dérivationnelles partielles sur la base des relations sémantiques qui les structurent	52
2.5	Représentation des paradigmes morphologiques dans le modèle <i>ParaDis</i>	53
2.6	Réseau sémantique structurant le paradigme en Table 2.9	59
3.1	Classe verbale DESTROY présentée dans <i>Verbnet</i> (Schuler, 2005)	76
3.2	Découpage de la famille de POT basé sur deux cadres distincts.	78
3.3	Définition et rôles principaux du cadre TUER présent dans <i>FrameNet</i> .	80
3.4	Rôles non-noyau (i.e. <i>non core frame elements</i>) qui apparaissent dans le cadre TUER dans <i>FrameNet</i> .	82
3.5	Extrait des types ontologiques utilisés dans <i>FrameNet</i>	83
3.6	Rôles noyau, relations inter-cadre et déclencheurs du cadre TUER	84

3.7	Exemple de relations reliant deux réseaux conceptuels distincts.	86
4.1	Paradigme dérivationnel structuré par des relations sémantiques (Bonami et Strnadová, 2019, p.172).	96
4.2	Graphe représentant le contenu sémantique de l’histoire en (42)	103
4.3	Représentation du contenu sémantique de l’histoire en (43)	106
4.4	Représentation du contenu de l’histoire en (44)	108
4.5	Graphe représentant le contenu de l’histoire en (45b)	111
4.6	Alignement de deux relations appartenant à deux histoires différentes	111
4.7	Alignement de plusieurs relations appartenant à des histoires différentes	113
4.8	Abstraction du contenu de deux histoires	114
4.9	SMB obtenu à partir du sous-ensemble $\{\text{POT.N, POTIER.N}\}$	116
4.10	SMB produit à partir de $\{\text{POT.N, POTIER.N}\}$ comportant un plan action et un plan activité	118
4.11	Représentation du contenu de l’histoire en (49)	121
4.12	SMB produit à partir de $\{\text{POT.N, REMPOTER.V, REMPOTEUR.N}\}$	122
4.13	Représentation du contenu de l’histoire en (50)	123
4.14	SMB produit à partir du contenu des histoires construites sur le sous-ensemble $\{\text{POT.N, POTIER.N, POTERIE.N, REMPOTER.V, REMPOTAGE.N}\}$	124
4.15	FMBs non connexes	126
4.16	FMB produit à partir de la famille de BOTTE structuré par une activité de fabrication de bottes ou d’autres chaussures.	129
4.17	FMB obtenu à partir de la famille de TOUR qui décrit l’activité de fabrication d’un artefact au tour et l’action qui l’instancie.	131
4.18	FMB obtenu à partir de la famille de FAÏENCE	132
4.19	FMB produit à partir de la famille de BOUTEILLE en (51d)	133
4.20	LMB décrivant une activité et une action de fabrication.	135
4.21	LMB décrivant le déplacement d’une entité concrète dans un récipient.	136

4.22 Schématisation de la procédure que nous avons adoptée pour obtenir des <i>lexical meaning bundles</i>	137
4.23 Exemple de projection d'un graphe représentant un LMB sur un graphe représentant une famille dérivationnelle	139
5.1 Représentation de la relation entre LOUVE et LOUVETEAU	150
5.2 FMB décrivant les relations sociales entre exemplaires d'un loup	152
5.3 LMB décrivant les relations sociales entre exemplaires d'une espèce animale sauvage	154
5.4 FMB produit à partir de la famille de FOURMI	157
5.5 LMB décrivant les relations entre un insecte, son nid et sa colonie	158
5.6 FMB décrivant la relation de type PROIE-PRÉDATEUR entre un fourmilier et une fourmi.	160
5.7 Obtention d'un LMB structuré par la relation entre une espèce animale prédatée et une espèce animale prédatrice.	162
5.8 FMB décrivant une activité de pêche à la crevette et une action qui l'instancie. .	164
5.9 LMB décrivant une activité de pêche et l'action qui l'instancie.	166
5.10 FMB décrivant l'activité de chasse au loup et une action qui l'instancie.	171
5.11 LMB décrivant une activité de chasse et l'action qui l'instancie.	172
5.12 LMB décrivant l'activité d'élimination d'animaux nuisibles.	175
5.13 FMB décrivant une activité d'élevage de crevettes	178
5.14 LMB structuré par une activité d'élevage en milieu aquatique	179
5.15 LMB décrivant une activité d'élevage d'animaux de ferme	181
5.16 LMB décrivant l'activité d'étude d'un animal	184
5.17 FMB décrivant la production des prunes.	189
5.18 LMB décrivant une activité de production fruiticole	190
5.19 FMB décrivant une activité de transport de bananes	192
5.20 LMB décrivant une activité de transport d'un bien commercialisé par un véhicule spécialisé.	194

5.21	FMB décrivant la préparation d'un bananier	196
5.22	LMB décrivant l'activité de préparation en cuisine.	198
5.23	Premier exemple de relation inter-LMB	200
5.24	Relations reliant un LMB décrivant la vie sociale des insectes et un LMB décrivant une activité de prédation de la part d'un animal	201
6.1	Descriptions des la relation masculin-féminin pour les participants humains des LMBs	210
6.2	Position des adjectifs relationnels dans les LMBs	213
6.3	Exemple d'intégration à un LMB d'un adverbe en <i>-ment</i> construit sur un Adj _{rel} par une relation périphérique de nature flexionnelle.	215
6.4	Dans un réseau action, le nom d'action est relié à son équivalent verbal au sommet qui représente l'action par une relation de transposition.	216
6.5	Représentation de l'itération dans un LMB	217

Liste des abréviations

ABL ablatif

ACC accusatif

A adjectif

Adj_{rel} adjectif relationnel

DAT datif

F féminin

FMB Family Meaning Bundle

GEN génitif

LMB Lexical Meaning Bundle

M masculin

MB Meaning Bundle

N nom

N_{fru} nom dénotant un fruit

N_{anim} nom dénotant un animal

NOM nominatif

PL pluriel

R adverbe

RCL Règle de Construction de Lexème

SG singulier

SMB Subset Meaning Bundle

V verbe

VOC vocatif

Introduction

La notion de *paradigme*, du grec ancien *παράδειγμα* (‘modèle’, ‘exemple’) est traditionnellement associée à la grammaire. Les apprenants de langues la rencontrent généralement lorsqu’ils étudient les conjugaisons verbales ou les déclinaisons nominales et adjectivales dans les manuels de grammaire. Les paradigmes sont conventionnellement représentés sous forme de tableau. Par exemple, le paradigme en Table 1 représente la première déclinaison des noms en latin. Dans ce paradigme, chacune des formes fléchies est associée à une combinaison de propriétés morphosyntaxiques : par exemple, *rosārŭm* est associé à la combinaison GEN.PL.

	SG	PL
NOM	<i>rosă</i>	<i>rosae</i>
GEN	<i>rosae</i>	<i>rosārŭm</i>
DAT	<i>rosae</i>	<i>rosīs</i>
ACC	<i>rosām</i>	<i>rosās</i>
VOC	<i>rosă</i>	<i>rosae</i>
ABL	<i>rosa</i>	<i>rosīs</i>

TABLE 1 – Exemple de paradigme en latin.

En morphologie, la nature paradigmaticque de la flexion est relativement consensuelle (Wurzel, 1989; Carstairs-McCarthy, 1991; Blevins, 2016; Bonami et Beniamine, 2016). Cette tendance s’est renforcée avec l’adoption de l’approche abstractive introduite par Blevins (2006) : cette approche s’attache à décrire les relations entre les formes de surface (i.e. des mots formes), plutôt qu’à fournir une grammaire capable de les générer (Beniamine, 2018, p.51). L’approche abstractive prend comme unité de base les mots formes et les relations qui les relient au sein d’un paradigme. Les mots formes sont vus comme des réalisations concrètes des *lexèmes*, à savoir des entités lexicales abstraites qui sont “stockées” (i.e. mémorisées) dans l’esprit du locuteur.

Une des propriétés principales associées aux paradigmes en flexion est l’interprédicibilité des éléments qu’ils contiennent. Cette interprédicibilité se situe à la fois sur le plan du contenu morphosyntaxique et sur le plan formel. Par exemple, si nous savons que dans le paradigme en Table 1 la forme *rosārŭm* exprime les propriétés morphosyntaxiques GEN.PL, nous pouvons prédire avec un certain degré de confiance que dans le même paradigme nous aurons une forme *rosām* qui exprime les propriétés morphosyntaxiques ACC.SG et vice versa. De la même manière, si nous savons que dans le paradigme du verbe LAYER en français le mot forme *laverai* exprime les propriétés morphosyntaxiques FUT.1SG, nous pouvons prédire que dans le même paradigme nous aurons une forme *laverons* qui exprime les propriétés FUT.1PL. La “force” de la prédiction dépend de la catégorie grammaticale du lexème, de sa classe flexionnelle et des propriétés morphosyntaxiques en question (Hathout et Namer, 2022, p.156).

L’adaptation des paradigmes à la dérivation a été longtemps rejetée à cause des nombreuses irrégularités qui caractérisent le lexique construit, par exemple la concurrence affixale entre *-age*, *-ion* et *-ment* en français ou entre *-ity* et *-ness* en anglais. Le lexique a été longtemps vu comme le lieu des idiosyncrasies : une unité lexicale peut intégrer le lexique d’une langue par construction morphologique, mais aussi par d’autres voies, comme les emprunts de mots à d’autres langues, l’attribution de nouveaux sens à des mots existants, etc. Toutefois, son organisation n’est pas anarchique : plusieurs études en sémantique lexicale soutiennent l’idée que le lexique est structuré par des relations récurrentes entre unités lexicales (ex. hyperonymie, synonymie, antonymie, etc.) et que ces relations permettent d’organiser le lexique sous forme d’un réseau (Cruse *et al.*, 1986; Murphy, 2003).

Dans le lexique construit, c’est-à-dire la partie du lexique qui est structurée par des relations morphologiques, plusieurs travaux proposent d’adapter la notion de paradigme pour décrire les régularités qui existent entre les relations dérivationnelles (Van Marle, 1985; Bauer, 1997; Štekauer, 2014; Stump, 2019; Hathout et Namer, 2019). Plusieurs travaux considèrent aussi que la notion de *famille dérivationnelle* occupe une place centrale dans l’organisation du lexique construit (Hathout, 2009; Haspelmath et Sims, 2013; Roché, 2017b). Une famille dérivationnelle peut être définie comme un ensemble de lexèmes reliés par des relations dérivationnelles, comme

en (1a). Cette notion couvre également des ensembles de lexèmes morphologiquement reliés dans lesquels il n’y a pas de base commune, comme en (1b).

- (1) a. LAVER, LAVABLE, LAVEUR, LAVEUSE, LAVERIE, LAVOIR, LAVETTE
b. AUDITEUR, AUDITRICE, AUDITION

Plusieurs travaux présentent les familles dérivationnelles comme l’équivalent des lexèmes en flexion (Haspelmath et Sims, 2013; Stump, 2019) et considèrent que les familles sont des paradigmes dérivationnels. La différence entre les deux types de paradigmes se situe sur le plan du contenu : les propriétés des mots formes qui organisent les paradigmes flexionnels sont morphosyntaxiques (voir Table 1), alors que dans un paradigme dérivationnel ces propriétés sont morphosémantiques. Par exemple, le nom LAVEUR dénote une personne qui réalise l’action de LAVER et RASOIR dénote un instrument que l’on utilise dans le processus dénoté par le verbe RASER.

Malgré les nombreuses références à la notion de famille dérivationnelle dans la littérature, la définition et la délimitation de cet objet restent un problème à résoudre (Hathout et Namer, 2019). Par exemple, l’extension d’une famille varie sensiblement en fonction des relations que l’on considère : est-ce que CHEVAL et ÉQUESTRE font partie de la même famille ? Qu’en est-il de PAPE et PONTIFICAT ? À cela s’ajoute le problème de l’interprédictibilité : si un paradigme contient des éléments interprédictibles qui assurent une certaine “cohérence interne”, comment se réalise cette cohérence interne dans les familles ? Par exemple, dans la famille de HERBE, il est difficile de soutenir que la présence d’HERBORISTE prédit sur le plan du sens et de la forme la présence de DÉSHERBER dans la même famille (et vice versa).

Bonami et Strnadová (2019) proposent de voir les paradigmes en dérivation comme des alignements de familles partielles contenant les mêmes relations. Ils adoptent une approche basée sur la sémantique et considèrent que les paradigmes sont des alignements de familles contenant les mêmes contrastes sémantiques. Un exemple est fourni dans la Table 2.

Bonami et Strnadová (2019) considèrent que les paradigmes sont des alignements de familles partielles sans pourtant préciser leurs frontières : par exemple, la famille de LAVER contient aussi

ACTION_V	ACTION_N	AGENT_N
LAVÉR	LAVAGE	LAVEUR
FORMER	FORMATION	FORMATEUR
GONFLER	GONFLEMENT	GONFLEUR

TABLE 2 – Exemple de paradigme dérivationnel proposé par Bonami et Strnadová (2019). La première ligne contient des propriétés sémantiques et catégorielles communes aux lexèmes de la même colonne (i.e. de la même série dérivationnelle)

LAVOIR et LAVERIE, mais les familles de FORMER et GONFLER ne contiennent pas de noms de lieu. Comment pouvons-nous déterminer le nombre et l’entité des cases disponibles dans un paradigme ? Et comment nous déterminons quelles (sous) familles peuvent y rentrer ? L’objectif principal de cette thèse est d’apporter une réponse à ces deux questions.

Objectifs et méthodologie

L’objectif principal de cette thèse est de proposer une méthode théorique et générale pour la délimitation des paradigmes dérivationnels en dérivation. L’analyse que nous réalisons porte sur le lexique construit du français en synchronie. Nous considérons que les paradigmes en dérivation sont formés par des familles dérivationnelles partielles qui présentent les mêmes relations de contraste, de la même manière que dans la proposition de Bonami et Strnadová (2019). Pour délimiter les familles et les organiser dans des paradigmes, nous adoptons une approche aux paradigmes dérivationnels guidée par la sémantique (Štekauer, 2014; Bonami et Strnadová, 2019; Fernández-Domínguez *et al.*, 2020). Selon cette approche, les relations sémantiques sont premières dans la structuration des paradigmes en dérivation.

Nous proposons de décrire les relations contenues dans les familles au moyen de réseaux conceptuels. Cela est en ligne avec des propositions récentes qui voient les familles dérivationnelles comme étant structurées par des réseaux conceptuels sous-jacents (Roché, 2017b). Nous identifions les relations et les limites des familles en utilisant des scénarios similaires à ceux qui sont proposés dans la Sémantique des Cadres (Fillmore *et al.*, 1976; Baker *et al.*, 1998; Petruck, 1996) et nous alignons à l’intérieur du même paradigme les familles contenant des lexèmes que

l'on peut inscrire dans le même scénario abstrait. Nous proposons d'abstraire ces scénarios à partir d'histoires. Les histoires sont dans cette thèse un outil heuristique que nous utilisons pour caractériser les relations entre les lexèmes contenus dans les familles dérivationnelles.

La méthode que nous proposons prend comme input une famille dérivationnelle "brute" comme celle en (2) et permet de la découper en familles partielles contenant des lexèmes interprédictibles sur le plan conceptuel. Dans ce cas, elle permet d'identifier les deux familles partielles en (3).

(2) {POT.N, POTIER.N, POTERIE.N, REMPOTER.V, REMPOTEUR.N}¹

(3) a. {POT.N, POTIER.N, POTERIE.N}

b. {POT.N, REMPOTER.V, REMPOTEUR.N}

Ensuite, notre méthode propose d'aligner les familles partielles obtenues avec ce découpage avec d'autres familles partielles qui contiennent les mêmes relations et qui peuvent être inscrites dans le même réseau conceptuel. Cela nous permet de former des paradigmes comme ceux qui sont proposés en Table 3 et Table 4. Le paradigme en Table 3 regroupe des familles partielles qui interviennent dans une activité de fabrication et le paradigme en Table 4 regroupe des familles qui interviennent dans une activité de déplacement dans un récipient. La première ligne de nos paradigmes décrit la structure sémantique du paradigme et comporte deux étiquettes : une étiquette relationnelle (ex. FABRICANT) et une étiquette ontologique (ex. PERSONNE), sauf pour les activités, qui sont le pivot sémantique du réseau. Dans nos paradigmes, nous admettons la possibilité qu'un lexème donné rentre dans plusieurs paradigmes (ex. POT qui apparaît dans les deux sous-ensembles en (3)) et qu'une famille donnée puisse ne pas "remplir" toutes les cases du paradigme (ex. pas de nom d'instrument pour rempoter). En outre, deux colonnes peuvent avoir la même étiquette ontologique (ex. ARTEFACT) : ce qui les discrimine est leur étiquette relationnelle (ex. OBJET FABRIQUÉ vs LIEU DE FABRICATION).

1. Dans cette thèse, nous notons les lexèmes en petites majuscules (ex.POT). En outre, nous notons les concepts et les catégories ontologiques et relationnelles en majuscule (ex. INSTRUMENT).

OBJET FABRIQUÉ ARTEFACT	FABRICANT PERSONNE	ACTIVITÉ.N	LIEU DE FABRICATION ARTEFACT
POT	POTIER	POTERIE	POTERIE
BOTTE	BOTTIER	BOTTERIE	BOTTERIE
BRIQUE	BRIQUETIER	BRIQUETERIE	BRIQUETERIE

TABLE 3 – Paradigmes regroupant un extrait des familles de lexèmes qui interviennent dans une activité de fabrication

DESTINATION RÉCIPIENT	ACTIVITÉ.V	AGENT DU DÉPLACEMENT. PERSONNE	INSTRUMENT ARTEFACT
POT	REMPOTER	REMPOTEUR	-
BOUEILLE	EMBOUEILLER	EMBOUEILLEUR	EMBOUEILLEUSE
CAISSE	ENCAISSER	ENCAISSEUR	-

TABLE 4 – Paradigme dérivationnel regroupant des familles de lexèmes qui interviennent dans une activité de déplacement dans un récipient.

À l'inverse de beaucoup de travaux en morphologie, qui prennent comme point de départ la forme (ex. un affixe) pour décrire ses fonctions sémantiques, nous prenons ici le chemin inverse en partant des concepts et des réseaux conceptuels pour arriver à leur réalisation morphologique. Nous faisons ce choix car la morphologie dérivationnelle est principalement une morphologie du lexique et l'une des fonctions du lexique est de permettre de nommer les choses, les événements et les propriétés. La dimension conceptuelle (ou onomasiologique) est partout sous-jacente lorsqu'on s'intéresse à la structuration du lexique (Roché, 2017a, p.2). Cette approche peut être résumée par la proposition de Bolozky (1999, p.7), qui considère que « *la formation lexicale est d'abord et avant tout basée sur la sémantique et axée sur les concepts.* »².

Nous illustrons notre méthode sur des familles dérivationnelles de nature sémantique variée comme des familles basées sur des noms d'artefact (ex. POT, POTIER, POTERIE), des familles basées sur noms de fruit (ex. POMME, POMMIER, POMMERAIE) et des familles basées sur des noms d'animaux (ex. LOUP, LOUVE, LOUVETIER).

2. « *lexical formation is first and foremost semantically based and concept driven* » dans le texte original.

Le projet Démonext

Cette thèse a été réalisée dans le cadre du projet *Démonext*³, qui a été financé par l'Agence Nationale de la Recherche (ANR), sous la référence ANR-17-CE23-0005. L'objectif général de ce projet était la création d'une base de données morphologiques dérivationnelles à large couverture du français : la base de données *Démonette*⁴.

Plan de la thèse

Nous présentons dans cette section l'organisation de la thèse.

Le **Chapitre 1** introduit les notions de *relation morphologique*, *famille dérivationnelle* et *série dérivationnelle*. Nous abordons la question du fonctionnement en synchronie des familles et de leur structuration en réseau. Nous y présentons quelques problèmes liés à la délimitation des familles, puis nous illustrons l'interface entre système conceptuel et lexique.

Dans le **Chapitre 2** nous présentons le cadre théorique de la morphologie dérivationnelle paradigmatique (Van Marle, 1985; Bauer, 1997; Bonami, 2017, parmi beaucoup d'autres). En particulier, nous nous concentrons sur la proposition selon laquelle les relations sémantiques structurent les paradigmes en dérivation. Nous présentons les modèles paradigmatiques de la dérivation proposés par Bonami et Strnadová (2019) et Hathout et Namer (2022) et des exemples de réseau qui ont été discutés dans des travaux de Roché (2017b) et Fradin (2021). Nous revenons sur le problème de la délimitation des paradigmes en dérivation, qui est le même que celui de la délimitation des familles cité dans le Chapitre 1.

Le **Chapitre 3** présente la notion de *rôle sémantique* et la théorie de la Sémantique des Cadres (Fillmore *et al.*, 1976; Petruck, 1996; Baker *et al.*, 1998). Les rôles sémantiques ont été employés dans plusieurs études en morphologie pour décrire les propriétés sémantiques des dérivés comme les noms déverbaux (Booij, 1986; Rosenberg, 2008; Salvadori et Huyghe, 2023). L'approche que nous adoptons dans cette thèse pour délimiter les familles paradigmatiques se

3. <https://www.demonext.xyz>

4. <https://demonette.fr>

base sur la notion de *cadre*, à savoir sur une situation cognitive dans laquelle interviennent des participants interprédictibles. Les cadres décrivent des scénarios qui sont indépendants de leur réalisation en discours et qui se situent sur le plan conceptuel.

Le **Chapitre 4** introduit la méthode générale que nous proposons pour le découpage des familles dérivationnelles en familles paradigmatiques alignables dans des paradigmes dérivationnels. Nous proposons de caractériser les relations entre lexèmes dans les familles au moyen d’histoires. Les histoires sont un outil heuristique que nous utilisons pour faire émerger les relations sémantiques récurrentes qui structurent les familles contenues dans le lexique construit. À partir des histoires, nous constituons des “paquets conceptuels” que nous appelons *meaning bundles* (MB). Les MBs contiennent les participants et les relations récurrentes des histoires et fonctionnent comme des sortes de scénarios abstraits, de la même manière que les cadres. La projection des MBs sur les familles nous permet de les découper et d’identifier les familles paradigmatiques, à savoir des familles qui peuvent être alignées dans des paradigmes.

Dans le **Chapitre 5** nous abordons plus en détail la question du rôle des propriétés ontologiques dans la structuration des paradigmes en dérivation. Nous illustrons notre proposition par des familles basées sur des noms de fruit et sur des noms d’animal et nous montrons la variété de paradigmes dérivationnels dans lesquels ces noms peuvent s’inscrire.

Dans le **Chapitre 6** nous discutons du statut de relations dérivationnelles “périphériques” par rapport aux MBs présentés au Chapitre 4. Il s’agit de relations comme celle qui connecte un adjectif relationnel et son nom de base (ex. JOURNALISTE.N - JOURNALISTIQUE.A), de la correspondance masculin-féminin chez les noms d’humain (ex. POTIER.N - POTIÈRE.N), ou encore le lien entre verbes et noms d’action exprimant le même prédicat (ex. LAYER.V -LAVAGE.N), entre adjectifs et adverbes dérivés en *-ment* ou enfin entre lexèmes ayant un sens prédicatif (ex. LAYER.V, LAVAGE.N) et leurs dérivés en *re-* (ex. RELAYER.V, RELAVAGE.N). Les relations périphériques complètent les réseaux conceptuels que nous proposons tout en étant indépendantes.

Publications et communications liées à la thèse

- Sanacore, D., Hathout, N., & Namer, F. (2023). A story-based approach to derivational paradigms. *International Symposium of Morphology 2023 (ISMo 2023)*.
- Sanacore, D., Hathout, N., & Namer, F. (2021). Frame-like structure for morphosemantic description. *Verbum (Presses Universitaires de Nancy)*, 43(1), 179-194.
- Sanacore, D., Hathout, N., & Namer, F. (2021). Scenarios and frames in derivation : a case study of derivational families based on animal names. *Resources and Tools for Derivational Morphology (DeriMo 2021)*, 114.
- Sanacore, D., Hathout, N., & Namer, F. (2020). Représentation sémantique des familles dérivationnelles au moyen de frames morphosémantiques. In *6e conférence conjointe Journées d'Études sur la Parole (JEP, 33e édition), Traitement Automatique des Langues Naturelles (TALN, 27e édition), Rencontre des Étudiants Chercheurs en Informatique pour le Traitement Automatique des Langues (RÉCITAL, 22e édition) (pp. 342-350). ATALA ; AFCP*.
- Sanacore, D., Hathout, N., & Namer, F. (2019, September). Semantic descriptions of French derivational relations in a families-and-paradigms framework. In *Proceedings of the Second International Workshop on Resources and Tools for Derivational Morphology (pp. 15-24)*.

Chapitre 1

Systeme conceptuel et organisation du lexique

Sommaire

1.1	Introduction	12
1.2	Relations dérivationnelles et lexique construit	13
1.3	Familles et réseaux	15
1.3.1	Délimitation des familles dérivationnelles	16
1.3.2	Nature et fonctionnement des familles	17
1.4	Séries dérivationnelles	20
1.5	Bilan intermédiaire	23
1.6	Interface entre lexique et système conceptuel	23
1.6.1	Types de relations conceptuelles	25
1.6.2	Relations lexicales	26
1.7	Compositionnalité du sens lexical	28
1.8	Représentation des réseaux conceptuels	31
1.9	Bilan	32

1.1 Introduction

Dans cette thèse nous nous intéressons en synchronie à la structuration du lexique construit français. Le lexique construit est la partie du lexique structurée par des relations morphologiques, à savoir des relations concomitantes de forme et de sens entre les mots. Dans ce chapitre, nous nous intéressons au lien entre organisation du lexique et système conceptuel.

En premier lieu, nous introduisons les notions de *relation morphologique*, de *famille dérivationnelle* et de *série dérivationnelle* : ces notions jouent un rôle central dans l’organisation du lexique construit (Hathout, 2009; Haspelmath et Sims, 2013; Fradin, 2018). Nous abordons en détail les problèmes de définition et la délimitation de la notion de famille : comme nous allons le voir, les lexèmes que l’on peut inclure dans une famille varient en fonction des relations que nous considérons (Hathout et Namer, 2019). Nous nous intéressons en particulier à la question de la structuration des familles : plusieurs travaux récents proposent de décrire la structure et le fonctionnement des familles au moyen de réseaux (Hathout et Namer, 2014; Roché, 2017b). Dans un réseau, les connexions entre les lexèmes de la famille sont multiples, contrairement à une structure arborescente qui contient des relations entre un input (la base) et un output (le dérivé). Pour décrire la nature des réseaux qui structurent les familles, des travaux récents proposent d’adopter une approche onomasiologique (Štekauer, 2014; Roché, 2017b). Selon cette approche, les familles sont structurées par des réseaux conceptuels sous-jacents qui structurent le lexique (construit) tout en étant indépendants de toute réalisation linguistique (Štekauer, 2014; Roché, 2017b; Fernández-Domínguez, 2019).

Le lien étroit entre système conceptuel et lexique est largement reconnu en sémantique lexicale (Cruse *et al.*, 1986; Pustejovsky, 1998; Talmy, 2000). Les concepts sont des unités de sens correspondant à des descriptions mentales des entités du monde. Plusieurs travaux en sémantique cognitive considèrent que les concepts ne sont pas isolés dans l’esprit du locuteur, mais s’organisent plutôt en “paquets” (Schank et Abelson, 1975; Fillmore *et al.*, 1976; Langacker, 1987; Croft *et al.*, 2004). Au sein du système conceptuel, les concepts sont reliés par des relations de différente nature. On distingue deux axes principaux : des relations de type taxinomique

(comme celles qui structurent les ontologies) et des relations de type engnomique, qui s'établissent entre concepts inscrits dans le même scénario cognitif et qui sont reliés par des relations de contiguïté.

Les relations lexicales reflètent en partie les relations existant au sein du système conceptuel (Cruse *et al.*, 1986). Après avoir introduit les relations lexicales “classiques” qui ont été largement discutées dans la littérature (ex. hypéronymie, hyponymie, méronymie) (Cruse *et al.*, 1986; Murphy, 2003), nous nous intéressons à l'usage des relations de contiguïté citées auparavant pour relier des unités lexicales sémantiquement proches comme COUTEAU et COUPER. Ces relations de contiguïté sont pertinentes pour la structuration des familles : par exemple, dans la famille d'un nom d'artefact comme COUTEAU, nous avons COUTELLIER et COUTELLERIE, qui dénotent respectivement une personne qui fabrique des couteaux et un lieu où les couteaux sont fabriqués.

En outre, nous abordons la question de la compositionnalité du sens lexical : plusieurs travaux en sémantique lexicale soutiennent l'idée que le sens lexical se compose d'un ensemble de propriétés sémantiques qui sont associées aux lexèmes (Cruse *et al.*, 1986; Pustejovsky, 1998). Ces propriétés sémantiques peuvent être organisées en “rubriques”, comme les *qualia* proposés par Pustejovsky (1998). Les *qualia* sont des différentes dimensions de sens associées à un lexème. Nous montrons la pertinence de ces dimensions de sens pour la structuration des familles dans le lexique construit.

Dans la dernière section de ce chapitre, nous introduisons le formalisme des *réseaux sémantiques* (Quillan, 1966; Sowa, 1992), que nous réadaptions dans cette thèse pour représenter la dimension sémantique des relations qui existent entre les membres des familles dérivationnelles et pour proposer une délimitation des familles sur base sémantique.

1.2 Relations dérivationnelles et lexique construit

Cette thèse se consacre à l'étude en synchronie de l'organisation du lexique construit français. Le lexique construit est la partie du lexique qui est structurée par des relations morphologiques,

à savoir des relations récurrentes de forme et de sens qui s'établissent entre les mots.

Définition 1

Relation morphologique

Une relation morphologique est une relation systématique de forme et de sens entre des mots (Haspelmath et Sims, 2013, p.2)

Les relations morphologiques comportent une dimension formelle et une dimension sémantique. Il existe une relation morphologique entre deux mots si : **(a)** il existe une relation de forme R_f entre ces deux mots ; **(b)** il existe une relation de contenu R_c entre les deux mots ; **(c)** il existe d'autres couples de mots reliés par la même relation de contenu R_c et la même relation de forme R_f (Bonami et Strnadová, 2019, p.3).

Les relations morphologiques auxquelles nous nous intéressons sont des relations dérivationnelles. Ces relations s'établissent entre des unités lexicales, à savoir des *lexèmes* (Matthews, 1974; Anderson et Anderson, 1992; Aronoff, 1993; Bonami *et al.*, 2018). Les lexèmes sont des entités abstraites qui sont concrètement réalisées par des mots formes. Par exemple, le lexème VIEUX⁵ est réalisé par quatre mots formes : *vieux*_{M.SG}, *vieille*_{F.SG}, *vieux*_{M.PL} et *vieilles*_{F.PL}.

Pour faire un exemple de relation dérivationnelle, nous considérons la relation entre le verbe LAYER et le nom déverbal LAVEUR. Cette relation comporte trois opérations distinctes :

- une opération affixale, à savoir une suffixation en *-eur* ;
- un changement de catégorie syntaxique, à savoir la création d'un nom à partir d'un verbe ;
- une opération sémantique qui permet de créer un nom dénotant une personne impliquée dans le prédicat dénoté par le verbe LAYER.

Une des fonctions principales de la morphologie dérivationnelle est celle de construire des unités lexicales (i.e. les lexèmes) qui peuvent intégrer le lexique. La dérivation est donc une

5. Comme annoncé dans l'introduction, dorénavant nous allons noter les lexèmes en petites majuscules, suivant en cela plusieurs auteurs (Matthews, 1974; Haspelmath et Sims, 2013; Bonami *et al.*, 2018, entre autres)

des principales sources d’enrichissement lexical⁶. Plusieurs auteurs soutiennent l’idée que l’organisation du lexique construit n’est pas “anarchique”, et que lorsqu’un lexème construit est inséré dans le lexique, il se place à la croisée entre une famille dérivationnelle et une série dérivationnelle. Ces deux notions sont présentées respectivement dans la Section 1.3 et la Section 1.4.

1.3 Familles et réseaux

Il est de plus en plus consensuel que la notion de *famille dérivationnelle* joue un rôle central dans la structuration du lexique construit (Hathout, 2009; Haspelmath et Sims, 2013; Roché, 2017b; Fradin, 2018, entre autres). Nous utilisons le terme “famille dérivationnelle” pour nous référer à un ensemble de lexèmes reliés par des relations dérivationnelles (Hathout et Namer, 2019).

Définition 2

Famille dérivationnelle

Une famille dérivationnelle est un ensemble de lexèmes reliés par des relations dérivationnelles.

Un exemple de famille dérivationnelle (partielle) est fourni en (4). Pour chaque lexème inscrit dans la famille, nous indiquons sa catégorie syntaxique.

- (4) LAVER.V, LAVEUR.N, LAVERIE.N, LAVABLE.A, LAVABILITÉ.N, LAVETTE.N

La définition de famille que nous venons de donner rend compte également d’ensembles de lexèmes qui n’ont pas de base commune mais qui pourtant sont morphologiquement apparentés. Par exemple, AUDITION et AUDITEUR sont morphologiquement apparentés même si leur base n’est pas attestée en français (pas de verbe *AUDIRE).

6. D’autre part, la dérivation peut également créer des occasionalismes (*nonce formations* en anglais) (Bauer, 1983), à savoir des mots complexes créés spontanément par les locuteurs pour satisfaire un besoin immédiat dans une situation de communication. Les occasionalismes sont créés sans l’intention d’imposer leur usage à une communauté linguistique (Dal et Namer, 2016, p.1).

Dans plusieurs travaux, la notion de famille est présentée comme étant l'équivalent des lexèmes en flexion : les relations flexionnelles relient les mots formes qui réalisent le même lexème et les relations dérivationnelles relient les membres d'une même famille dérivationnelle (Haspelmath et Sims, 2013; Hathout et Namer, 2019). En revanche, Hathout et Namer (2019) soulignent que, contrairement aux lexèmes en flexion, les familles sont des ensembles ouverts, dont la délimitation reste un problème à résoudre. Nous approfondissons ce point dans la section suivante.

1.3.1 Délimitation des familles dérivationnelles

Le problème principal lié à la notion de famille est la délimitation de ses "frontières". Déterminer quand une famille est complète ou partielle et quels lexèmes peuvent y rentrer reste une question à clarifier. Par exemple, Hathout et Namer (2019, p.6) montrent que l'extension d'une famille varie sensiblement en fonction des relations que l'on considère. On peut considérer qu'une famille dérivationnelle inclut uniquement des lexèmes reliés par des relations morphologiques établies dans le lexique en synchronie (ex. LAYER - LAYER), ou bien inclure aussi des emprunts (ex. CHEVAL - ÉQUESTRE) ou des lexèmes uniquement reliés par le sens (ex. TOMBER - CHUTE).

Au sein des familles dérivationnelles, Hathout (2009) propose une distinction entre deux sous-types de familles dérivationnelles : les *familles dérivationnelles morphologiques* (ou *familles morphologiques*) et les *familles dérivationnelles lexicales* (ou *familles lexicales*). Les familles dérivationnelles morphologiques contiennent des lexèmes qui partagent des propriétés de forme et de sens (Hathout, 2009, p.37), comme la famille proposée en (5a). D'autre part, les *familles dérivationnelles lexicales* rassemblent des lexèmes appartenant à plusieurs familles dérivationnelles morphologiques complémentaires qui sont apparentées au niveau du sens. Par exemple, la famille lexicale en (5b) contient également des lexèmes qui sont reliés à CHEVAL uniquement au niveau du sens et qui sont construits sur des thèmes savants, comme ÉQUESTRE et HIPPIQUE, ou empruntés à d'autres langues comme CAVALIER (emprunté à l'italien) (Fradin, 2018, p.3).

- (5) a. CHEVAL.N, CHEVALIER.N, CHEVALERIE.N, CHEVALIN.A, CHEVALET.N, CHEVAUCHER.V
 b. CHEVAL.N, CHEVALIER.N, CAVALIER.N, ÉQUIN.A, ÉQUESTRE.A, HIPPIQUE.A

Un des objectifs de cette thèse est de proposer une méthode pour délimiter les familles dérivationnelles et déterminer leurs frontières. Pour pouvoir apporter une réponse à la question de la délimitation des familles, il est nécessaire de s’interroger sur comment les familles sont structurées. Nous abordons cette question dans la section suivante.

1.3.2 Nature et fonctionnement des familles

Nous venons de voir que la notion de famille dérivationnelle, tout en étant largement adoptée dans la littérature, reste problématique à délimiter. Pour pouvoir apporter une première réponse à la question de la délimitation des familles, nous nous interrogeons sur leur structure. Une première approche adoptée dans la littérature propose de représenter les familles comme étant des arborescences. Ces arborescences sont structurées par des relations entre un input (i.e. la base) et un output (i.e. le dérivé). La Figure 1.1 montre l’exemple utilisé par Stump (2019, p.3) pour représenter la famille de **INVENTER**. Une représentation de ce type adopte une approche “génétique” à la description de la structuration des familles : on part de l’élément plus simple pour aller vers le plus complexe.

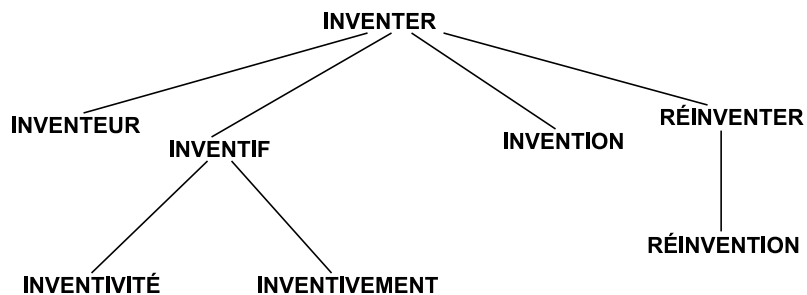


FIGURE 1.1 – Représentation sous forme d’arborescence d’une famille dérivationnelle partielle (Stump, 2019, p.3).

En revanche, cette approche pose des problèmes pour traiter des cas où l’orientation d’une relation morphologique n’est pas claire, comme pour le couple MÉDECIN -MÉDECINE. En outre,

Roché (2017b) souligne que l'intérêt de situer les lexèmes construits dans les familles est de voir comment les unités lexicales "fonctionnent" les unes par rapport aux autres et comment la famille existante conditionne les nouvelles créations destinées à l'intégrer (Roché et Plénat, 2014, 2016). Une représentation en arborescence ne rend pas compte des relations existant entre des membres d'une famille créés à partir de la même base qui partagent des propriétés formelles et sémantiques. Par exemple, dans la famille en Figure 1.1 on pourrait considérer qu'il existe bien une relation entre INVENTEUR et INVENTION. Ces deux lexèmes partagent des propriétés formelles et sur le plan sémantique : un inventeur est nommé ainsi car il a produit une ou plusieurs inventions et chaque invention a son inventeur. De la même manière, dans le triplet en (6) ENSEIGNEMENT et ENSEIGNANT sont tous les deux construits à partir d'ENSEIGNER, mais ils sont également motivés l'un par rapport à l'autre : l'activité professionnelle de l'enseignant est l'enseignement et l'enseignement est le métier de l'enseignant. La relation sémantique est présente dans le fonctionnement de la famille et marche dans les deux sens.

(6) ENSEIGNER.V, ENSEIGNEMENT.N, ENSEIGNANT.N

Pour rendre compte du fonctionnement en synchronie des familles dans le lexique, plusieurs travaux récents adoptent une approche "fonctionnelle" et non génétique à la description des familles. Ils proposent de représenter la structure des familles comme étant un réseau, plutôt qu'une arborescence qui va du plus simple au plus complexe (Hathout et Namer, 2014; Roché, 2017b). Contrairement à une structure en arborescence, où l'on a des relations de type BASE → DÉRIVÉ, dans un réseau les connexions sont multiples et fonctionnent dans les deux sens. La figure 1.2 représente sous forme de graphe les relations entre les lexèmes en (6).

Dans la proposition de Roché (2017b) pour décrire la structure et le fonctionnement des familles, la dimension sémantique joue un rôle fondamental. Comme d'autres travaux récents (Štekauer, 2014; Antoniova et Štekauer, 2016; Fernández-Domínguez, 2019), il applique une approche onomasiologique aux familles dérivationnelles et considère que les familles sont structurées par des réseaux de nature conceptuelle. En morphologie, traditionnellement, on adopte une approche sémasiologique : on part par la forme (mots, affixes) pour décrire le sens. Pour sa

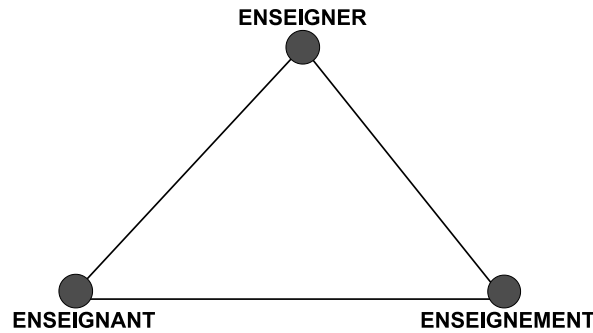


FIGURE 1.2 – Représentation de la famille partielle de ENSEIGNER sous forme de graphe.

part, Roché (2017b) propose de décrire la structure des familles en partant par le sens, à savoir par le réseau conceptuel sous-jacent, pour arriver à la famille.

Les réseaux conceptuels qui structurent les familles sont indépendants de toute réalisation linguistique et se situent uniquement sur le plan sémantique (Štekauer, 2014; Roché, 2017b; Fradin, 2021; Hathout et Namer, 2022). Par exemple, d'un point de vue conceptuel, dans la famille en Figure 1.2 on a une activité et une personne qui s'y consacre et ces concepts existent indépendamment de leur réalisation dans le lexique et dans la même famille.

Roché (2017b) considère que les familles dérivationnelles plus simples sont structurées par un seul réseau, alors que les familles plus complexes sont structurées par plusieurs réseaux qui peuvent éventuellement se chevaucher. Par exemple, dans la famille en (7a), le sens de JOUR est bien présent dans JOURNAL, mais n'intervient pas dans le sens lexical de JOURNALISME et JOURNALISTE. Sur le plan sémantique, il n'y a pas de raison d'inclure JOUR dans le même réseau conceptuel que JOURNALISME et JOURNALISTE. Pour sa part, la famille en (7b) contient POTIER et REMPOTER : ces deux lexèmes ne sont pas à priori sémantiquement connexes, car un potier ne s'occupe pas de rempoter des plantes dans son métier et donc on pourrait considérer qu'ils n'appartiennent pas au même réseau conceptuel. En revanche, POTERIE est bien motivé par rapport à POTIER, car la poterie est le métier du potier et le potier fait de la poterie.

- (7) a. JOUR.N, JOURNAL.N, JOURNALISME.N, JOURNALISTE.N
 b. POT.N, POTIER.N, POTERIE.N, REMPOTER.V, REMPOTEUR.N

Il est donc possible que deux réseaux coexistent dans la famille en (7b) et que ces deux réseaux aient en commun la présence de POT : par rapport à POTIER le pot est l'objet construit, alors que par rapport à REMPOTER un pot est une sorte d'instrument qui intervient dans le processus. Cette question est approfondie en Section 1.7 où nous introduisons la question de la compositionnalité du sens lexical et sa pertinence pour la description des relations entre les lexèmes des familles.

Nous considérons qu'une approche guidée par la sémantique pour la description des familles dérivationnelles permet non seulement de rendre compte de leur fonctionnement en synchronie, mais peut éventuellement apporter une réponse au problème de la délimitation des familles que nous avons cité précédemment. Si les familles se structurent en réseaux et si ces réseaux sont de nature conceptuelle, c'est peut-être sur une base sémantique et sur la base de scénarios conceptuels que l'on peut délimiter les familles dérivationnelles et identifier les différents réseaux qui les structurent. Ce point est approfondi au Chapitre 2, où nous aborderons également la notion de *paradigme* et le problème de définition et délimitation des paradigmes en dérivation.

1.4 Séries dérivationnelles

Une autre notion centrale dans la structuration du lexique construit est la notion de série. Un lexème construit, lorsqu'il est intégré dans le lexique, se place à la croisée entre une famille dérivationnelle, que nous venons de présenter, et ce qu'on appelle une *série dérivationnelle* (Roché, 2011). Une série se compose de lexèmes qui sont construits au moyen du même procédé morphologique. Un exemple est présenté en (8) : cette série contient des noms déverbaux en *-eur*.

(8) LAVEUR.N, JOUEUR.N, VOLEUR.N, RONFLEUR.N

Les familles dérivationnelles et les séries dérivationnelles sont en relation par le fait que chaque lexème contenu dans une famille appartient à une série distincte. Par exemple, si nous reprenons la famille partielle de LAYER en (9), nous pouvons établir que cette famille contient

un lexème appartenant à la série des noms déverbaux en *-age* (10a), un lexème appartenant à la série des noms déverbaux en *-eur* (10b) et un lexème appartenant à la série des noms déverbaux en *-oir* (10c).

(9) LAVER.V, LAVEUR.N, LAVAGE.N, LAVOIR.N

(10) a. LAVAGE.N, RASAGE.N, RECYCLAGE.N

b. LAVEUR.N, VOLEUR.N, CONSTRUCTEUR.N

c. LAVOIR.N, FUMOIR.N, ABATTOIR.N

De la même manière que pour les familles, Hathout (2009, p.38) propose une distinction au sein des séries dérivationnelles. Il en distingue deux types : les *séries dérivationnelles morphologiques* (ou *séries morphologiques*) et les *séries dérivationnelles lexicales* (ou *séries lexicales*). Les *séries morphologiques* ne contiennent que de lexèmes qui ont été construits au moyen du même procédé dérivationnel. Par exemple, la série en (11) contient des noms d'action en *-ion* construits à partir de verbe. Pour leur part, les *séries dérivationnelles lexicales* généralisent le contenu de plusieurs séries morphologiques : par exemple, la série dérivationnelle lexicale en (12) contient des lexèmes comme *LOCOMOTION* et *CONFECTION* qui ne peuvent pas être insérés dans la série dérivationnelle morphologique en (11) car ils ne sont pas en relation avec un verbe en synchronie. Une série dérivationnelle lexicale réunit plusieurs séries dérivationnelles morphologiques uniquement sur la base de leur forme ou de leur sens (Hathout, 2011, p.263).

(11) COMPILATION.N, PRODUCTION.N, CONSTRUCTION.N

(12) COMPILATION.N, PRODUCTION.N, CONSTRUCTION.N, CONFECTION.N, LOCOMOTION.N

La Table 1.1 propose un récapitulatif des types de familles et de séries morphologiques et lexicales que nous avons introduits jusqu'ici dans les Sections 1.3.1 et 1.4.

En outre, Hathout *et al.* (2020) font une distinction supplémentaire au sein des séries dérivationnelles en distinguant les *séries de lexèmes* et les *séries de relations*. Les séries de lexèmes sont des ensembles de lexèmes qui présentent les mêmes contrastes de forme et de sens avec les

	MORPHOLOGIQUE	LEXICALE
FAMILLE	CHEVAL, CHEVALIER, CHEVALERIE...	CHEVAL, CHEVALIER, CHEVALERIE, HIPPIQUE, HIPPODROME ...
SÉRIE	COMPILATION, PRODUCTION, CONSTRUCTION...	COMPILATION, PRODUCTION, CONSTRUCTION, CONFECTION, LOCOMOTION ...

TABLE 1.1 – Exemples des différents types de familles et de séries proposés par Hathout (2009)

membres de leur famille. Un exemple est la série des noms d'action en *-age* (13). Les séries de relations, pour leur part, sont des ensembles de couples de lexèmes qui présentent les mêmes contrastes de forme et de sens (14). Une série de relations est donc un cas particulier de série de lexèmes.

(13) LAVAGE.N, SOUDAGE.N, CASSAGE.N, ÉCLAIRAGE.N ...

(14) LAVER.V - LAVAGE.N, CASSER.V - CASSAGE.N, ÉCLAIRER.V - ÉCLAIRAGE.N,
SOUDER.V - SOUDAGE.N ...

Comme pour les familles, les éléments (ou plutôt les couples) qui peuvent rentrer dans une série de relations dépendent des relations que l'on considère : la série en (14) contient des couples de lexèmes qui sont reliés par la même relation de forme et la même relation de sens (i.e. relation entre verbe et nom d'action en *-age*). En revanche, si on prend comme point de départ la dimension sémantique des relations morphologiques pour décrire la structuration des familles, alors les couples en (15) peuvent faire partie de la même série de relations, car ils contiennent tous un verbe et un nom dénotant ce même prédicat. Ce point est approfondi au Chapitre 2, où nous abordons la question du rôle des familles et de séries dans la structuration des paradigmes dérivationnels.

(15) LAVER.V - LAVAGE.N, FORMER.V - FORMATION.N, GONFLER.V - GONFLEMENT.N, VOLER.V - VOL.N

1.5 Bilan intermédiaire

Pour récapituler, jusqu'ici nous avons présenté les notions de famille et de série dérivationnelles. Nous avons mis en évidence le problème de délimitation de la notion de famille et nous avons abordé la question de leur délimitation. Nous avons vu que des travaux récents proposent de représenter les familles comme étant des réseaux et nous avons vu que certains auteurs considèrent que ces réseaux sont de nature conceptuelle. Dans cette thèse, nous adoptons la proposition de Roché (2017a) et nous considérons que les familles et la structure du lexique construit s'appuient sur des réseaux conceptuels sous-jacents. Pour décrire la structuration des familles, nous allons prendre comme point de départ la dimension sémantique des relations qu'elles contiennent. Nous considérons que cette approche guidée par la sémantique peut aussi être adoptée pour résoudre le problème de délimitation des familles, qui est un des objectifs de cette thèse. La question de la délimitation des familles et du rôle des séries dans la structuration du lexique est reprise au Chapitre 2, où nous introduisons la question de l'usage de la notion de *paradigme*.

Dans la suite de ce chapitre, nous abordons plus en détail la notion de *concept* et l'organisation des concepts dans le système conceptuel. En particulier, nous nous concentrons sur le lien entre système conceptuel et relations lexicales. En outre, nous introduisons la question de la compositionnalité du sens lexical en montrant sa pertinence pour la structuration des familles.

1.6 Interface entre lexique et système conceptuel

L'idée que le lexique et le système conceptuel sont étroitement reliés est désormais largement admise en sémantique lexicale (Lyons, 1977; Cruse *et al.*, 1986; Pustejovsky, 1998; Koch, 2005; Murphy, 2003, entre autres). Par exemple, selon Cruse *et al.* (1986), les concepts jouent un rôle intermédiaire entre la réalité extralinguistique et les unités lexicales.

Dans le domaine de la sémantique cognitive, les concepts sont considérés comme étant les unités de base de la cognition humaine (Langacker, 1987; Talmy, 2000). Löbner (2013, p.19)

utilise ce terme pour se référer à la description mentale associée à une entité. Il précise qu'un concept ne correspond pas forcément à une description visuelle : par exemple, les référents des concepts BRUIT ou ERREUR ne décrivent pas d'entités visibles⁷. En outre, le contenu d'un concept associé à une entité ne se limite pas à la description de ses propriétés physiques. Par exemple, la description du concept CHIEN ne se limite pas à l'apparence physique des entités appartenant à la catégorie des chiens, mais comporte tout ce que nous connaissons sur cette catégorie de référents : leur comportement, leur rôle dans notre société (animal de compagnie, chien guide, chien de sauvetage, chien de chasse), etc.

Un concept donné peut être réalisé sous forme d'unité lexicale dans un procédé dit de *lexicalisation* (Levin et Hovav, 2005). Cependant, la relation entre concepts et unités lexicales n'est pas biunivoque : on peut avoir une unité lexicale réalisant plusieurs concepts (i.e. polysémie) et un ensemble d'unités lexicales qui réalise le même concept (i.e. synonymie). En outre, on peut avoir dans le système conceptuel des concepts qui ne sont pas réalisés dans le lexique d'une langue. Sur cette base, nous pouvons définir un concept comme la description mentale d'un référent potentiel associé à un lexème, comme le fait Löbner (2013).

Définition 3

Concept

Un concept est une description mentale associée à un référent potentiel d'un lexème.

Les concepts sont des unités mentales indépendantes de toute réalisation linguistique et qui se fondent sur notre connaissance du monde. En revanche, la connaissance du monde varie d'une culture à une autre : pour cette raison, plusieurs auteurs considèrent les concepts comme étant basés sur des aspects culturels (Fillmore *et al.*, 1976; Lakoff, 1988; Croft *et al.*, 2004).

Dans le domaine de la sémantique cognitive, il est de plus en plus consensuel que les concepts n'existent pas de manière isolée dans l'esprit du locuteur mais sont organisés en "paquets" (Schank et Abelson, 1975; Fillmore *et al.*, 1976; Langacker, 1987; Koch, 2000; Talmy, 2000; Croft *et al.*, 2004). L'exemple utilisé pour illustrer cette idée est celui du concept RESTAURANT :

7. Comme annoncé dans l'introduction, nous notons les concepts en majuscule, suivant en cela Croft *et al.* (2004), pour les distinguer des lexèmes. Par exemple, on a CHIEN (concept) et chien (lexème).

pour caractériser ce concept, nous devons forcément faire appel à d'autres concepts qui lui sont étroitement reliés comme CLIENT, SERVEUR, MENU, CUISINE ou COMPTOIR. Cela signifie que le concept RESTAURANT, pour être caractérisé, ne peut pas être isolé des autres concepts qui font partie de son même "scénario". La notion de *cadre* (Fillmore *et al.*, 1976), que nous approfondissons au Chapitre 3, se base sur cette hypothèse.

Le système conceptuel est donc structuré par des relations entre les concepts et un concept est défini en prenant en compte les autres concepts auxquels il est relié. Dans la section suivante, nous présentons les deux types principaux de relations conceptuelles parmi celles identifiées par Koch (2005), les *relations taxinomiques* et les *relations engynomiques*. Ces deux types de relations définissent deux dimensions du système conceptuel.

1.6.1 Types de relations conceptuelles

Au sein du système conceptuel, Koch (1999, 2005) identifie deux dimensions principales sur lesquelles les relations conceptuelles se situent, à savoir les relations taxinomiques et les relations engynomiques. Sa classification des relations conceptuelles est conçue pour rendre compte de la structuration du lexique dans une perspective typologique, en prenant en compte des données lexicales provenant de langues variées.

- Les *relations taxinomiques* s'inscrivent dans les hiérarchies conceptuelles. Cette dimension relie des concepts subordonnés partageant des propriétés sémantiques comme CHAT, TIGRE, PUMA à des concepts superordonnés par lesquels ces propriétés sémantiques sont héritées, comme FÉLIN (Blank, 2003, p.10) ;
- Les *relations engynomiques* sont des relations de contiguïté dans le cadre d'une situation cognitive qui se base sur l'expérience du monde réel. La dimension engynomique concerne des concepts inscrits dans le même scénario (ou cadre, voir Chapitre 3), comme des relations partie/tout, cause/conséquence, producteur/produit, etc. Par exemple, un concept comme CHAT est relié par des relations de contiguïté à des concepts comme GRIFFE, LITIÈRE ou RONRONNEMENT.

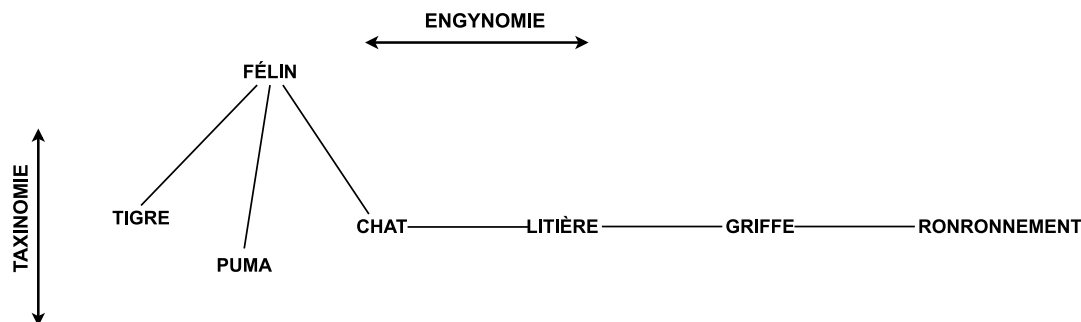


FIGURE 1.3 – Concepts associés au concept CHAT sur le plan taxinomique et sur le plan engynomique.

Ces deux types de relations sont schématisés en Figure 1.3 : les relations taxinomiques structurent le système conceptuel de manière “verticale”, alors que les relations engynomiques réunissent des concepts au moyen de relations de contiguïté. Comme nous le verrons plus en détail dans la section suivante, l’organisation lexicale reflète en partie les relations qui s’établissent au niveau conceptuel sur le plan taxinomique et sur le plan engynomique.

1.6.2 Relations lexicales

De nombreux travaux se sont intéressés à la structure du lexique et à la description des relations sémantiques existantes entre les unités lexicales (Lyons, 1977; Cruse *et al.*, 1986; Pustejovsky, 1998; Murphy, 2003). Ces travaux ont mis en évidence l’existence de relations sémantiques régulières entre unités lexicales et voient ces régularités comme le reflet de la présence de relations conceptuelles qui sont pertinentes pour la structuration du lexique. Plus précisément, Cruse observe que certaines relations de sens connectent les unités lexicales au moyen de patrons sémantiques récurrents : « *certaines relations de sens entre les mots reviennent systématiquement, liant les mots dans des schémas sémantiques récurrents.* » (Cruse, 2002, p.83)⁸. Pour sa part, Nutter (1989, p.2) considère que les relations lexicales fournissent un moyen formel pour exprimer des relations conceptuelles. Nous présentons ci-dessous les principales relations lexicales qui ont été discutées dans la littérature.

⁸. « *Certain meaning relationships between words crop up again and again, binding the words into recurrent semantic patterns* » dans le texte original.

- *synonymie* : deux lexèmes sont synonymes lorsqu'ils réalisent le même concept dans le lexicque (ex. VOITURE -AUTOMOBILE). Deux unités lexicales sont synonymes si elles peuvent se substituer l'une à l'autre dans un contexte linguistique sans en altérer le sens (Miller *et al.*, 1990, p.7).
- *antonymie* : il s'agit d'une relation qui marque un contraste fort au niveau du sens, comme pour les adjectifs JEUNE - VIEUX. Une sous-spécification de cette relation a été proposée par Cruse *et al.* (1986) ;
- *hyperonymie* : cette relation relie un lexème renvoyant à un concept superordonné à un autre lexème renvoyant à un concept subordonné qui en hérite des propriétés sémantiques (ex. FÉLIN - CHAT) ;
- *hyponymie* : cette relation relie un lexème renvoyant à un concept subordonné à un autre lexème renvoyant à un concept superordonné, dont il hérite des propriétés sémantiques (ex. CHAT - FÉLIN) ;
- *méronymie* : cette relation regroupe une gamme de relations partie-tout (ex. DOSSIER - CHAISE). La méronymie a été sous-spécifiée par Winston *et al.* (1987), qui en définit six types différents ;
- *troponymie* : cette relation connecte un couple de verbes dont l'un exprime une manière particulière d'effectuer l'action exprimée par l'autre (ex. MURMURER-PARLER).

En revenant à notre problème de délimitation et de description des familles, la question qui se pose est si les relations lexicales classiques que nous venons d'introduire permettent de décrire et de délimiter les familles. Nous considérons qu'elles ne sont pas suffisantes pour notre objectif, car elles ne sont pas utilisables pour décrire des relations morphologiques comme MÉDECIN-MÉDECINE, LAYER-LAVOIR ou POT-POTIER que nous avons citées en Section 1.3.

Des relations supplémentaires aux relations classiques citées auparavant sont adoptées dans des ressources lexicales comme *WordNet* (Miller *et al.*, 1990; Fellbaum, 2010), qui utilise des *rôles sémantiques* (Gruber, 1965; Fillmore, 1967) pour rendre compte de certaines relations morphosémantiques. Par exemple, la relation de PEINTRE par rapport à PEINDRE est marquée comme étant de type AGENT et la relation entre SLEEPER 'couchette' par rapport à SLEEP

‘dormir’ est marquée comme étant de type LIEU⁹. L’inconvénient de l’inventaire des relations morphosémantiques fournies par *Wordnet* est que sa couverture est très limitée.

D’autre part, des nombreux travaux ont proposé d’étendre l’inventaire des relations lexicales classiques en extrayant des relations lexicales à partir de textes (Nutter, 1989; Morris et Hirst, 2004; Nastase *et al.*, 2022, entre autres). Par exemple, certains inventaires comme celui de Cassidy (2000) ou celui de Nutter (1989) contiennent des centaines de relations lexicales extraites à partir de textes. Certaines relations inscrites dans ces inventaires pourraient être pertinentes pour nos besoins de description des familles : par exemple, dans l’inventaire proposé par Nutter (1989), qui à son tour regroupe les résultats d’une dizaine d’autres travaux, on trouve des relations comme FEMELLE DE qui pourrait être utilisée pour relier LIONNE à LION, PETIT DE que l’on pourrait utiliser pour relier LOUVETEAU à LOUVE, OBJET - USAGE qui peut relier RASOIR et RASER, ACTE - INSTRUMENT qui peut relier RASER et RASOIR et ACTE - LIEU qui peut relier FUMER et FUMOIR. D’autre part, ces inventaires sont très vastes et hétérogènes et ne sont pas spécialement “calibrés” pour la morphologie dérivationnelle ni pour le français. Pour ces raisons, nous n’exploitons pas directement un de ces inventaires, mais certaines relations que nous allons retrouver dans la suite de cette thèse reprennent les exemples de relations que nous venons de fournir.

Dans la section suivante nous abordons un autre aspect qui caractérise les unités lexicales et que nous jugeons comme pertinent pour la structure des familles, à savoir la compositionnalité du sens lexical.

1.7 Compositionnalité du sens lexical

Comme nous l’avons vu en Section 1.6, un concept est généralement inscrit dans plusieurs réseaux conceptuels. Par exemple, le concept CHAT peut être caractérisé par rapport à l’apparence physique des chats, à leur comportement, à leur nature de félins, à leurs proies, etc. Du

9. Nous approfondissons la question de l’usage des rôles sémantiques pour décrire les relations morphologiques au Chapitre 3. Nous notons les rôles sémantiques en majuscule, comme les concepts (ex. AGENT).

côté du lexique, les unités lexicales présentent une pluralité de propriétés sémantiques, par rapport auxquelles ces unités lexicales peuvent être mises en relation avec d'autres unités lexicales. Dans cette perspective, le Lexique Génératif (LG) proposé par Pustejovsky (1998), adopte une description compositionnelle du sens lexical des lexèmes pour rendre compte de la structuration sémantique du lexique. Dans le LG, le sens d'une unité lexicale L est décrit au moyen des rubriques suivantes.

- *structure événementielle* : définition du type d'événement dénoté par l'unité lexicale : état, processus ou transition.
- *structure argumentale* : spécification du nombre et du type sémantique des arguments de l'unité lexicale. Elle spécifie quels arguments sont obligatoires et lesquels sont optionnels.
- *structure des qualia* : spécification des attributs essentiels de l'objet dénoté par L .

Dans cette section, nous nous concentrons sur la structure des *qualia*. Cette structure se compose de quatre sous-rubriques, ou facettes suivant la terminologie de Cruse *et al.* (1986). Chacune de ces sous-rubriques remplit un rôle particulier dans la description du sens d'une unité lexicale L (Pustejovsky, 1998; Busa *et al.*, 2001; Pustejovsky et Jezek, 2016). Nous les présentons ci-dessous.

- *rôle formel* : décrit des relations de nature taxonomique en indiquant la façon dont l'entité référencée par L se définit par rapport au reste du monde. La relation d'hyponymie s'inscrit dans cette dimension ;
- *rôle constitutif* : énumère les éléments constitutifs du référent de L et exprime des relations de type partie - tout (Namer, 2009, p.273). Les relations méronymiques rentrent dans cette rubrique ;
- *rôle agentif* : énumère les conditions qui doivent être remplies pour permettre l'existence du référent de L . Par exemple, pour les artefacts, cette rubrique permet de préciser les événements nécessaires à la création du référent de L . Elle exprime des relations de type *créé par* (Pustejovsky et Jezek, 2016, p.7). Pour les noms agentifs permanents comme FUMEUR ou VIOLONISTE, le rôle indique l'origine de cette pérennité comme l'habitude pour FUMEUR et l'aptitude pour VIOLONISTE (Busa, 1997) ;

- *rôle télélique* : il précise la finalité du référent de L et exprime des relations de type *utilisé-pour*. Par exemple, la relation entre COUTEAU et COUPER s’inscrit dans cette rubrique.

Pour reprendre l’exemple de POT, la structure des *qualia* permet de décrire son contenu sémantique par rapport à la nature artefactuelle de son référent, par rapport à son origine, par rapport à la manière dont il est fabriqué, à sa fonction (i.e. contenir des choses) et aux parties qui les constituent. Sur la base du contenu des *qualia* pour le lexème POT nous pouvons établir les relations lexicales suivantes :

- Par rapport au *quale* formel, POT peut être mis en relation avec un hyperonyme comme ARTEFACT ;
- Par rapport au *quale* constitutif, POT peut être mis en relation avec ARGILE, car l’argile est utilisée comme matériau pour faire des pots ;
- Par rapport au *quale* agentif, POT peut être mis en relation avec POTIER, car un pot est généralement créé par un potier. La relation entre POT et POTERIE se situe également sur cette dimension ;
- Par rapport au *quale* télélique, POT peut être mis en relation avec REMPOTER, car ce verbe dénote un accomplissement pour lequel l’utilisation d’un POT est nécessaire à la réalisation.

Dans la famille de POT on a donc des relations sémantiques qui sont pertinentes pour des facettes distinctes du sens lexical de POT : sa nature et son usage. Ces deux facettes correspondent à deux scénarios distincts où le pot est impliqué et donc à deux réseaux conceptuels distincts.

Comme nous allons le voir dans la suite de cette thèse, les familles contiennent de manière régulière des relations qui se situent au niveau de *qualia* distincts d’un même lexème. Dans d’autres cas, les propriétés sémantiques pertinentes pour les relations morphologiques sont encore plus spécifiques : par exemple dans la famille de BANANE nous trouvons BANANIER, qui dénote à la fois un arbre de bananes, un gâteau dont l’ingrédient principal est la banane et un cargo spécialisé dans le transport des bananes. Comme nous allons le voir au Chapitre 5, ce n’est pas la nature de fruit de BANANE qui est pertinente pour sa relation avec BANANIER

(cargos), mais plutôt le fait que les bananes ont besoin d'être transportées dans des conditions de température qui permettent de les conserver.

Pour récapituler, les relations dérivationnelles qui structurent les familles peuvent se situer au niveau de propriétés sémantiques distinctes d'un même lexème (ex. nature du pot vs usage du pot). Cela signifie que dans la famille de POT coexistent au moins deux réseaux décrivant deux scénarios dans lesquels POT joue un rôle différent.

1.8 Représentation des réseaux conceptuels

Comme nous l'avons montré précédemment dans ce chapitre, nous considérons qu'une approche guidée par la sémantique peut nous permettre à la fois de caractériser les relations entre lexèmes appartenant aux familles dérivationnelles et d'apporter une réponse au problème de la délimitation des familles. Pour représenter les réseaux sémantiques dont nous nous servons, nous réadaptions le formalisme des réseaux sémantiques (Quillan, 1966; Sowa, 2014), que nous présentons dans cette section.

Les réseaux sémantiques sont utilisés pour représenter toute connaissance qui peut être décrite par le langage naturel (Lehmann, 1992, p.2). Ils se basent sur la notion de *graphe*. Un graphe se compose de sommets et d'arcs. Dans les réseaux sémantiques, les sommets représentent des concepts (entités, actions, propriétés) et les arcs représentent des relations entre concepts. Dans un réseau sémantique, les arcs et les sommets sont généralement étiquetés : l'étiquette associée à un sommet identifie le concept et l'étiquette associée à l'arc identifie le type de relation. En Figure 1.4 nous montrons un exemple de représentation de connaissance sous forme de réseau sémantique. Le réseau à gauche contient à la fois des relations orientées de type taxinomique (i.e. EST-UN) et des réseaux qui se situent sur le plan engnomique (i.e. CHASSE). À droite, nous donnons un aperçu de comment une structure basée sur des sommets et des arcs étiquetés peut rendre compte du contenu sémantique des relations que l'on a dans les familles. La représentation à droite en Figure 1.4 est une version "préliminaire" des réseaux sémantiques que nous proposons au Chapitre 4 pour caractériser les relations entre lexèmes des familles. Le

réseau à droite comporte des relations qui ne sont pas orientées et qui marchent dans les deux sens : JOURNALISME, JOURNALISTE et JOURNAL sont tous sémantiquement motivés l'un par rapport à l'autre.

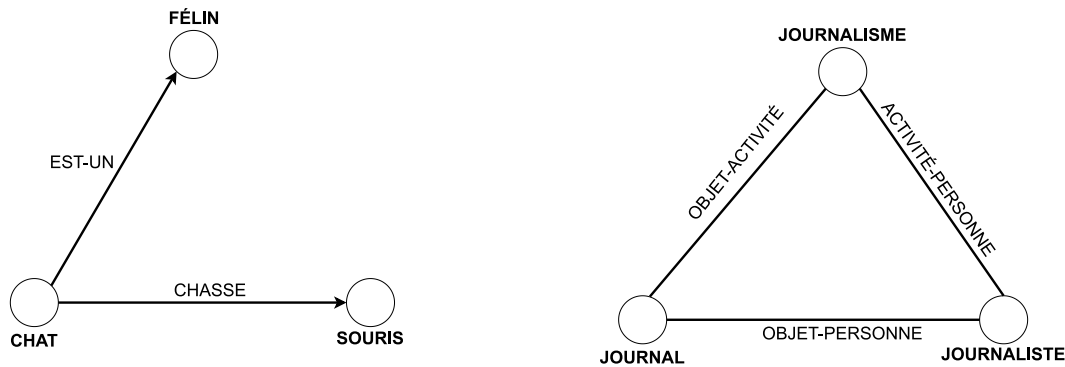


FIGURE 1.4 – Exemple de représentation de connaissance sous forme de réseau sémantique. Dans un réseau sémantique, les sommets et les arcs sont étiquetés. Les relations peuvent être orientées (un CHAT est un FÉLIN, mais pas l'inverse) ou peuvent marcher dans les deux sens.

Il existe des nombreux types de réseaux sémantiques qui ont été adoptés dans la littérature pour des fonctions variées comme la représentation du discours ou la structuration d'ontologies (Sowa, 2014, pour un aperçu). Dans cette thèse, nous nous limitons à nous servir de schémas contenant des arcs étiquetés (orientés ou non orientés selon les besoins) et des sommets étiquetés. En outre, de manière similaire à ce qui est fait dans les réseaux sémantiques, nous allons nous servir de sommets de formes différentes pour distinguer entités, événements et propriétés, comme nous le montrons au Chapitre 4.

1.9 Bilan

Dans ce chapitre, nous avons abordé la question de la structuration du lexique construit et de l'interface entre structuration du lexique et système conceptuel. Nous avons commencé par décrire le rôle des notions de famille et de série dans l'organisation du lexique construit. En outre, nous avons montré que la notion de famille dérivationnelle, qui est souvent évoquée dans la littérature pour décrire la structuration du lexique construit, reste problématique à cerner et

nous avons annoncé qu'un des objectifs principaux de cette thèse est d'apporter une réponse à ce problème.

Pour répondre à la question de la délimitation des familles, nous adoptons une approche guidée par la sémantique qui nous permet de décrire leur structuration. Pour décrire cette approche, nous avons introduit en particulier la proposition de Roché (2017a), qui considère que les familles sont structurées par une ou plusieurs réseaux sémantiques/conceptuels qui les structurent.

Nous avons introduit dans la discussion sur la structuration du lexique les notions de concept et les deux principales dimensions sur lesquelles les relations entre concepts se situent : la dimension taxonomique et la dimension engnominique. Nous avons ensuite vu que les relations lexicales reflètent en partie les relations existant au niveau conceptuel et nous avons parcouru les principales relations lexicales qui ont été discutées dans la littérature.

Nous avons constaté que les inventaires de relations lexicales existant à la fois dans la littérature et dans des ressources lexicales de grande taille comme *WordNet* ne sont pas directement exploitables : certains, comme celui de *Wordnet*, sont trop limités et d'autres, comme celui de Nutter (1989), contiennent des relations qui pourraient être réadaptées, mais ces inventaires ne sont pas conçus spécialement pour la morphologie dérivationnelle et contiennent des relations de nature très hétérogène. Pour cette raison, nous n'exploitons pas directement un jeu de relations lexicales déjà existant pour décrire les relations inscrites dans les familles, mais certaines relations que nous allons utiliser dans cette thèse reprennent les relations présentées dans ce chapitre et les rôles sémantiques que nous allons présenter au Chapitre 3.

Nous avons ensuite discuté la question de la compositionnalité du sens lexical et de comment les relations morphologiques contenues dans une famille peuvent être pertinentes pour des *qualia* ou des propriétés sémantiques distinctes qui caractérisent le sens lexical d'un lexème.

Pour conclure, nous avons présenté rapidement le formalisme que nous réadaptions dans cette thèse pour représenter les réseaux qui structurent les familles dérivationnelles : il s'agit des réseaux sémantiques (Quillan, 1966; Sowa, 2014).

Dans le Chapitre 2 nous abordons une autre question centrale qui caractérise l'organisation

du lexique construit et les relations morphologiques, à savoir la régularité des relations contenues dans les familles. Dans cette perspective, nous abordons la notion de paradigme dérivationnel et nous discutons comment l'approche basée sur la sémantique que nous avons introduite dans ce chapitre permet de caractériser la notion de paradigme dérivationnel.

Chapitre 2

Une approche sémantique aux paradigmes dérivationnels

Sommaire

2.1	Introduction	36
2.2	La notion de paradigme	38
2.2.1	Les paradigmes et leur usage en flexion	38
2.2.2	Adapter les paradigmes à la dérivation	40
2.2.3	À quoi ressemblent les paradigmes en dérivation ?	41
2.2.4	Apport des paradigmes à l'analyse dérivationnelle	47
2.3	Exemples de modèles paradigmatiques	50
2.3.1	Les systèmes paradigmatiques de Bonami et Strnadová (2019)	50
2.3.2	Le modèle paradigmatique ParaDis	52
2.4	Variété sémantique des paradigmes en dérivation	55
2.4.1	Paradigme basé sur des noms de fruits	55
2.4.2	Paradigme basé sur les ethniques	57
2.4.3	Réseau action	59
2.4.4	Réseau activité	62
2.5	Bilan	64

BILAN DU CHAPITRE 1 : La notion de famille dérivationnelle joue un rôle central dans l’organisation du lexique construit. Cependant, la délimitation de cette notion reste un problème à résoudre dans la littérature (Hathout et Namer, 2019). Des travaux récents considèrent que les familles sont structurées par des réseaux conceptuels (Štekauer, 2014; Roché, 2017b). Au sein du système conceptuel, les concepts sont reliés par des relations de nature variée (taxinomiques, de contiguïté) et s’organisent en réseau. L’interface entre lexique et système conceptuel a été discutée dans plusieurs travaux en sémantique lexicale : les relations lexicales reflètent en partie les réseaux contenus dans le système conceptuel (Cruse *et al.*, 1986; Murphy, 2003).

2.1 Introduction

Dans le Chapitre 1 nous avons vu qu’une relation dérivationnelle est une relation récurrente de forme et de sens qui s’établit entre des lexèmes appartenant à une même famille dérivationnelle. Dans ce chapitre, nous abordons la notion de paradigme dérivationnel, qui peut être définie de manière générale comme un patron récurrent au sein du lexique construit.

Si en morphologie flexionnelle l’organisation paradigmatique est relativement consensuelle (Carstairs-McCarthy, 1994; Wunderlich et Fabri, 1995; Ackerman *et al.*, 2009; Blevins, 2016; Bonami et Beniamine, 2016), l’usage des paradigmes en dérivation est encore débattu dans la littérature. En effet, la dérivation a été longtemps considérée comme irrégulière et caractérisée par une faible prédictibilité (Dressler, 1989; Plank, 1994). Cela est dû à des phénomènes de “non canonicité” qui caractérisent le lexique construit, comme la compétition affixale, les décalages forme - sens, la défectivité, etc.

D’autre part, au cours des dernières décennies plusieurs auteurs ont proposé d’adopter une analyse paradigmatique des relations dérivationnelles et du lexique construit (Van Marle, 1985; Bauer, 1997; Hathout et Namer, 2019; Ruz *et al.*, 2022, entre autres). Plusieurs conférences (p.ex. *ParadigMo 2017* et *ParadigMo 2021*) et ouvrages thématiques (*Lingue e Linguaggio XVII(2)*, 2018, et *Morphology 29(2)*, 2019) ont été consacrés à cette question. L’approche paradigmatique

de la dérivation vise à rendre compte des régularités qui caractérisent le lexique construit et permet d’envisager un traitement unifié de la flexion et la dérivation comme faisant partie d’une seule morphologie (Stump, 1991; Bauer, 1997; Štekauer, 2014; Bonami et Strnadová, 2016; Boyé et Schalchli, 2019).

Comme nous allons le voir, les notions de famille et de série que nous avons introduites dans le Chapitre 1 interviennent dans la constitution des paradigmes dérivationnels. Par exemple, des travaux récents considèrent que les paradigmes dérivationnels sont des alignements de familles partielles qui contiennent les mêmes relations (Bonami et Strnadová, 2019; Hathout et Namer, 2022). De manière similaire à ce que nous avons fait dans le Chapitre 1 pour décrire la structuration des familles, dans ce chapitre nous nous concentrons sur une approche guidée par la sémantique de la structuration des paradigmes dérivationnels (Štekauer, 2014; Antoniova et Štekauer, 2016; Bonami et Strnadová, 2019). Selon cette approche, deux familles peuvent être inscrites dans le même paradigme dérivationnel si elles sont structurées par les mêmes relations sémantiques. Dans cette perspective, les relations formelles entre les lexèmes sont secondaires.

D’autre part, le problème de la délimitation des familles dérivationnelles que nous avons cité dans le Chapitre 1 reste et “engendre” un problème de délimitation des paradigmes dérivationnels. Dans la deuxième partie du chapitre, nous discutons deux modèles paradigmatiques récents, à savoir les systèmes paradigmatiques de Bonami et Strnadová (2019) et *ParaDis*, qui a été proposé par Hathout et Namer (2022), pour voir si les modèles paradigmatiques de la dérivation existants permettent d’apporter une réponse à cette question. En d’autres termes, nous voulons voir si ces modèles permettent de “découper” les familles dérivationnelles et, par conséquent, les paradigmes formés en alignant les familles dérivationnelles qui contiennent les mêmes relations.

Dans la dernière partie de ce chapitre, nous approfondissons la question de la nature sémantique des paradigmes dérivationnels et nous présentons des exemples de paradigmes de nature sémantique variée qui ont été discutés dans des travaux récents (Boyé et Schalchli, 2017; Roché, 2017b; Fradin, 2020a).

2.2 La notion de paradigme

Avant d’aborder la notion de *paradigme* telle qu’elle est utilisée pour décrire le lexique construit, nous introduisons son usage pour la description des relations flexionnelles, auquel les paradigmes sont plus traditionnellement associés. Ensuite, en Section 2.2.2, nous introduisons la question de l’adaptation des paradigmes pour décrire la structuration du lexique construit.

2.2.1 Les paradigmes et leur usage en flexion

La nature paradigmatique des relations flexionnelles est aujourd’hui relativement consensuelle (Carstairs-McCarthy, 1994; Wunderlich et Fabri, 1995; Stump, 1991; Blevins, 2006; Ackerman *et al.*, 2009). Il existe deux sens principaux du terme *paradigme* dans l’organisation des relations flexionnelles. Ces deux sens sont présentés comme suit par Carstairs-McCarthy (1991, p.639).

- (16) a. *paradigme*₁ : l’ensemble des combinaisons de propriétés morphosyntaxiques (i.e. l’ensemble des cases) réalisées par les formes fléchies des lexèmes pour une catégorie morphosyntaxique donnée dans une langue donnée.
- b. *paradigme*₂ : l’ensemble des réalisations flexionnelles qui expriment le *paradigme*₁ pour un lexème donné dans une langue donnée.

Les paradigmes sont traditionnellement représentés dans un format tabulaire. Un exemple de *paradigme*₁ est proposé en Table 2.1. Ce paradigme se compose de quatre cases. Il s’agit d’un paradigme abstrait : le contenu des cases est déterminé par toutes les combinaisons possibles de valeurs morphosyntaxiques pour un adjectif en français (i.e. masculin singulier, féminin singulier, masculin pluriel, féminin pluriel).

M.SG	M.PL
F.SG	F.PL

TABLE 2.1 – Paradigme abstrait d’un adjectif en français, qui correspond au *paradigme*₁ proposé par Carstairs-McCarthy (1991).

Pour sa part, l'exemple de paradigme₂ en Table 2.2 est la réalisation flexionnelle de l'adjectif NATIONAL, où chaque mot forme réalise une combinaison de valeurs morphosyntaxiques contenue dans le paradigme abstrait en Table 2.1.

<i>national</i>	<i>nationaux</i>
<i>nationale</i>	<i>nationales</i>

TABLE 2.2 – Paradigme flexionnel de l'adjectif NATIONAL en français, qui correspond au paradigme₂ proposé par Carstairs-McCarthy (1991).

Un aspect central qui caractérise les paradigmes flexionnels est l'interprédictibilité des éléments qu'ils contiennent, à la fois sur le plan des propriétés morphosyntaxiques et de leur réalisation formelle (Wurzel, 1989; Boyé, 2011; Bonami, 2014; Beniamine, 2018; Pellegrini, 2020). Par exemple, si nous savons que le paradigme de l'adjectif NATIONAL en français contient un mot forme *national* qui exprime la combinaison de valeurs morphosyntaxiques M.SG, on peut prédire avec un certain degré de confiance que dans le même paradigme nous allons trouver un mot forme *nationaux* qui exprime les valeurs morphosyntaxiques M.PL. Le degré de prédictibilité varie en fonction de la case du paradigme à laquelle on s'intéresse, de la classe flexionnelle à laquelle appartient le lexème concerné et de sa catégorie grammaticale (Hathout et Namer, 2022, p.156).

Dans un paradigme flexionnel, tous les mots formes sont reliés par des relations morphologiques récurrentes. Cela est schématisé dans l'exemple en Figure 2.1 emprunté à Bonami et Strnadová (2019, p.5) : on a un paradigme abstrait d'un adjectif en français (équivalent à celui en Table 2.1) et on a trois paradigmes flexionnels concrets représentant respectivement les lexèmes ÉGAL, PETIT et VIEUX. Les mots formes de chaque paradigme sont reliés par des relations binaires et qui marchent dans les deux sens : chaque paradigme en Figure 2.1 contient six relations. Ces relations sont récurrentes sur le plan du contenu morphosyntaxique. Dans le modèle proposé par Bonami et Strnadová (2019), les relations flexionnelles sont alignées sur la base de leur contenu morphosyntaxique et leur réalisation formelle est secondaire. Cet aspect relatif au modèle paradigmatique de Bonami et Strnadová (2019) est approfondi en Section 2.3.1.

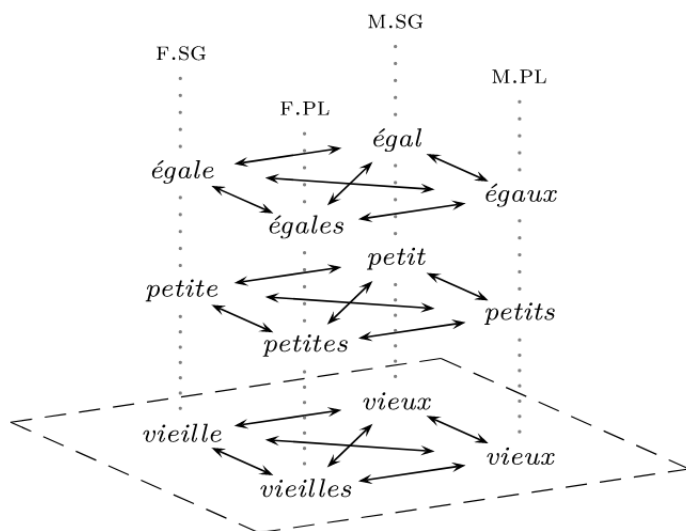


FIGURE 2.1 – Exemple de trois paradigmes concrets contenant les mêmes relations sur le plan des valeurs morphosyntaxiques. Ces trois paradigmes réalisent le même paradigme abstrait (i.e. paradigme₁) (Bonami et Strnadová, 2019, p.6)

Pour récapituler, les paradigmes en flexion sont des objets contenant des éléments interprédictibles sur le plan de leur contenu morphosyntaxique. Les relations flexionnelles qui structurent les paradigmes sont des relations récurrentes. Le contenu récurrent de ces relations peut être représenté sous forme de paradigme abstrait (qui correspond à la notion paradigme₁ définie par Carstairs-McCarthy (1991)).

2.2.2 Adapter les paradigmes à la dérivation

Dans l'introduction de son article sur les paradigmes dérivationnels, Štekauer (2014) souligne que « Une partie inhérente de presque toutes les discussions sur les différences entre la flexion et la dérivation était l'affirmation selon laquelle la morphologie flexionnelle est paradigmatique alors que la morphologie dérivationnelle ne l'est pas. » (Štekauer, 2014, p.354)¹⁰. Cette affirmation tient compte des nombreuses irrégularités qui caractérisent la dérivation, par exemple de la

10. Texte original : « An inherent part of almost all discussions of the differences between inflection and derivation used to be the claim that inflectional morphology is paradigmatic while derivational morphology is not. »

compétition affixale des suffixes *-age*, *-ion* et *-ment* en français (Dal *et al.*, 2018; Fradin, 2019; Missud et Villoing, 2020; Huyghe et Varvara, 2023).

D'autre part, des travaux réalisés dans les dernières décennies ont montré que plusieurs raisons qui motivent l'usage des paradigmes en flexion valent aussi en dérivation (Bauer, 1997; Stump, 2005; Spencer, 2013; Bonami et Strnadová, 2016). En premier lieu, la flexion et la dérivation se servent des mêmes opérations morphologiques (affixation, reduplication, etc.). En outre, les lexèmes construits que l'on trouve dans les familles expriment des catégories conceptuelles sémantiques récurrentes dans le lexique comme AGENT, PATIENT, INSTRUMENT, LIEU, ITÉRATION ou RÉSULTAT (ce point est approfondi au Chapitre 3), de manière similaire à ce qui font les mots formes dans les paradigmes flexionnels, qui pour leur part réalisent des valeurs morphosyntaxiques récurrentes (voir Section 2.2.1).

Concernant le problème des nombreuses irrégularités de la dérivation, ces travaux soutiennent que la flexion comporte aussi des phénomènes d'irrégularité (supplétion thématique, syncrétisme, etc.), même s'ils y sont moins fréquents (Corbett, 2010; Walther, 2013), mais que cela n'empêche pas de représenter les relations flexionnelles comme étant paradigmatiques.

La différence principale entre paradigmes flexionnels et paradigmes dérivationnels est liée à la nature obligatoire de la flexion, alors que la réalisation d'une catégorie conceptuelle dans un paradigme dérivationnel dépend des besoins de la langue (Štekauer, 2014, p.357). En outre, les paradigmes dérivationnels présentent un problème qui *a priori* ne se pose pas en flexion : si les paradigmes sont représentables au moyen de tables, quelles sont les cellules de ces tables ? Cette question revient au problème de la délimitation des paradigmes en dérivation, que nous approfondissons dans la section suivante.

2.2.3 À quoi ressemblent les paradigmes en dérivation ?

Dans les travaux qui proposent une approche paradigmatique des relations dérivationnelles, le terme *paradigme* a été utilisé dans des sens variés. Plusieurs auteurs utilisent ce terme pour se référer aux familles dérivationnelles, que nous avons présentées dans le Chapitre 1 (Bauer, 1983;

Antoniova et Štekauer, 2016; Stump, 2019, entre autres). D'autres travaux utilisent le terme *réseau dérivationnel* (*derivational network* en anglais) (Körtvélyessy *et al.*, 2020, par exemple) pour se référer à des familles dérivationnelles partielles qui réalisent dans le lexique un réseau conceptuel donné. Dans d'autres travaux encore, les paradigmes dérivationnels sont présentés comme des empilements de familles dérivationnelles (partielles) contenant les mêmes relations (Bonami et Strnadová, 2019; Hathout et Namer, 2022) et qui peuvent être alignées sur la base de ces relations.

En premier lieu, nous testons la première approche que nous venons de citer en considérant un exemple de famille dérivationnelle en français en Figure 2.2. Comme nous l'avons souligné en Section 2.2.1, une des propriétés principales des paradigmes en morphologie est l'interprédictibilité des éléments qu'ils contiennent sur le plan du contenu (nous laissons de côté l'aspect formel pour l'instant). Sur le plan du contenu (morphosyntaxique en flexion et morphosémantique en dérivation), la présence d'un élément x dans un paradigme prédit la présence d'un élément y dans le même paradigme et vice versa.

En revanche, il est difficile de soutenir que dans la famille en Figure 2.2 la présence de HERBORISTE prédit sur le plan sémantique la présence de DÉSHERBANT, ou que la présence de HERBACÉE peut être prédite au niveau sémantique par la présence de DÉSHERBER dans la même famille. Ces lexèmes dénotent des référents qui ne sont pas sémantiquement proches : un herboriste vend des herbes médicinales, alors qu'un désherbant est un produit qui détruit des mauvaises herbes.

D'autre part, la famille en Figure 2.2 contient effectivement des sous-ensembles de lexèmes qui sont sémantiquement proches et que l'on peut considérer comme étant interprédictibles sur le plan sémantique. Par exemple, HERBE, HERBORISTE et HERBORISTERIE sont sémantiquement interprédictibles : HERBORISTE dénote une personne qui vend des herbes et HERBORISTERIE dénote la boutique de l'herboriste et le métier de l'herboriste. Un autre sous-ensemble de lexèmes interprédictibles comprend HERBE, DÉSHERBER, DÉSHERBAGE et DÉSHERBANT : DÉSHERBER et DÉSHERBAGE dénotent l'action d'éliminer des mauvaises herbes et DÉSHERBANT dénote un produit utilisé dans ce processus.

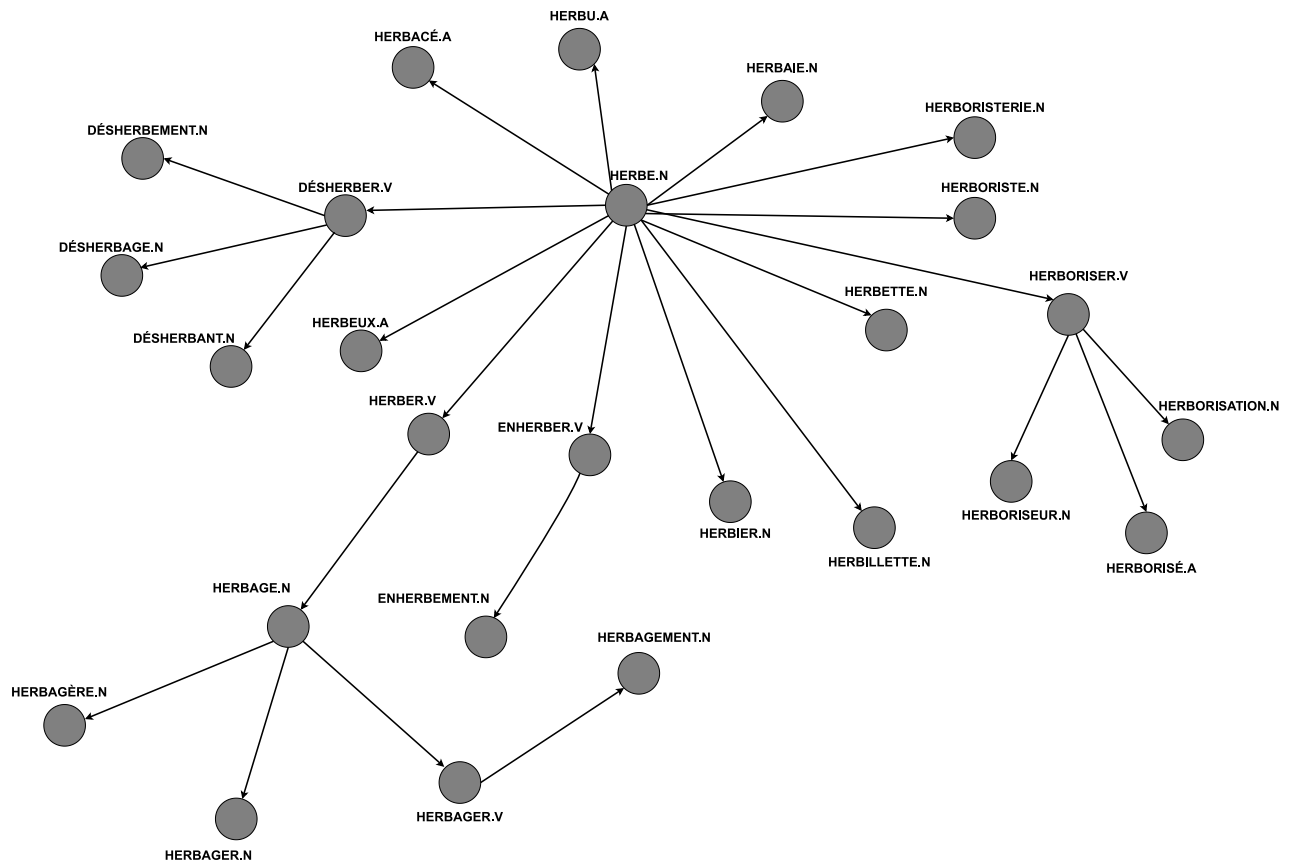


FIGURE 2.2 – Représentation sous-forme de réseau de la famille de HERBE. Les arcs orientés représentent des relations base \rightarrow dérivé.

Sur cette base, nous considérons que : **(a)** les familles dérivationnelles “complètes” ne peuvent pas constituer des paradigmes car les lexèmes qu’elles contiennent ne sont pas assez interprédictibles sur le plan sémantique ; **(b)** en revanche, dans les familles dérivationnelles, nous pouvons identifier des sous-ensembles qui contiennent des lexèmes interprédictibles sur le plan sémantique et qui peuvent former des paradigmes.

En outre, les relations qui structurent les paradigmes sont des relations récurrentes. Par conséquent, nous nous attendons que les relations qui relient les sous-ensembles de lexèmes interprédictibles dans la famille de la Figure 2.2 soient présentes aussi dans d’autres familles dans le lexique construit.

Ce constat introduit une deuxième approche des paradigmes dérivationnels : selon cette approche, les paradigmes dérivationnels sont des empilements de familles qui présentent les mêmes relations. Par exemple, le paradigme fourni en Table 2.3 contient quatre familles (partielles) en français qui présentent des régularités formelles et sémantiques (Hathout et Namer, 2022, p.158).

LAVÉR	LAVAGE	LAVEUR	LAVEUSE	LAVABLE
CASSER	CASSAGE	CASSEUR	CASSEUSE	CASSABLE
ÉCLAIRER	ÉCLAIRAGE	ÉCLAIREUR	ÉCLAIREUSE	ÉCLAIRABLE
SOUDER	SOUDAGE	SOUDEUR	SOUDEUSE	SOUDABLE

TABLE 2.3 – Exemple de paradigme dérivationnel en français. Ce paradigme contient quatre familles contenant les mêmes relations morphologiques et sémantiques.

Les régularités à la fois formelles, catégorielles et sémantiques qui caractérisent le paradigme en Table 2.3 peuvent être représentées au moyen du paradigme abstrait (paradigme₁ dans la terminologie de Carstairs) proposé en Table 2.4. Dans ce paradigme abstrait, la caractérisation sémantique des membres du paradigme est faite au moyen de catégories sémantiques comme ACTION, AGENT et POSSIBILITÉ. Le contenu des colonnes dans la Table 2.4 correspond aux propriétés formelles, catégorielles et sémantiques communes aux lexèmes inscrits dans chaque colonne en Table 2.3. Chaque colonne de la Table 2.3 correspond à une série dérivationnelle.

X	Xage	Xeur	Xeuse	Xable
V	N	N	N	A
‘action’	‘action’	‘agent’	‘agent’	‘possibilité’

TABLE 2.4 – Paradigme abstrait qui décrit les régularités dans le paradigme en Table 2.3

Une autre question qui se pose pour la caractérisation des paradigmes dérivationnels concerne le type de régularités à considérer pour constituer un paradigme : on pourrait se demander si un paradigme dérivationnel doit nécessairement contenir des familles alignées à la fois sur le plan formel et sur le plan sémantique ou si un de ces deux aspects est premier par rapport à l’autre.

Par exemple, Antoniova et Štekauer (2016) et Bonami et Strnadová (2019) adoptent une approche qui considère les relations sémantiques comme étant premières pour la structuration

des paradigmes. Cette approche permet d’inscrire dans le même paradigme des dérivés sémantiquement équivalents mais qui sont réalisés avec des moyens formels différents, comme dans l’exemple en Table 2.5, qui est adapté par Bonami et Strnadová (2019). Ce paradigme contient trois familles partielles structurées par les relations existant entre un verbe dénotant une action, un nom dénotant le même processus et un nom dénotant une personne qui intervient dans ce processus en tant qu’agent. Une approche aux paradigmes dérivationnels guidée par la sémantique est également partagée par Bauer (2019) et Fernández-Domínguez *et al.* (2020, p.3), qui considèrent que le sens plutôt que la forme, est le facteur décisif pour les paradigmes en dérivation.

LAVÉ	LAVEUR	LAVAGE
FORMER	FORMATEUR	FORMATION
LANCER	LANCEUR	LANCEMENT
VOLER	VOLEUR	VOL

TABLE 2.5 – Exemple de paradigme dérivationnel structuré par des relations sémantiques adapté par Bonami et Strnadová (2019, p.6)

Dans cette thèse, nous rejoignons les propositions de ces auteurs et nous considérons que les relations sémantiques sont premières pour la structuration des paradigmes en dérivation et que l’aspect formel est secondaire. Des familles (partielles) sont alignables dans un paradigme si leurs lexèmes réalisent les mêmes “catégories conceptuelles” et sont reliés par les mêmes relations sémantiques. Pour le dire autrement, nous considérons que des familles dérivationnelles peuvent être alignées dans un paradigme si elles sont structurées par le même réseau sémantique sous-jacent. Sur cette base, la définition de paradigme dérivationnel que nous utilisons dans cette thèse est la suivante :

Définition 4

Paradigme dérivationnel

Un paradigme dérivationnel est un empilement de familles dérivationnelles (partielles) contenant les mêmes relations sémantiques.

Une autre question qui se pose dans la définition des paradigmes dérivationnels est le problème de la défektivité : jusqu’ici nous avons présenté des paradigmes dérivationnels sous forme de tables où toutes les cases étaient remplies (p.ex. voir Table 2.3), mais que faire lorsqu’une famille présente des trous ? Cette situation est exemplifiée par la Table en 2.6 : la famille dans la première ligne contient SCIERIE, qui dénote une usine où on utilise des scies pour découper le bois ou le marbre, alors que les deux autres familles ne contiennent pas un nom de lieu.

SCIE	SCIER	SCIAGE	SCIEUR	SCIERIE
FRAISE	FRAISER	FRAISAGE	FRAISEUR	-
TOUR	TOURNER	TOURNAGE	TOURNEUR	-

TABLE 2.6 – Exemple d’alignement de familles dérivationnelles “incomplètes”

À ce propos, Štekauer (2014) considère que le fait que certaines familles contiennent des trous n’empêche pas d’avoir des paradigmes. Il considère que la prédictibilité et la régularité des paradigmes en dérivation doivent être vues en termes de disponibilité conceptuelle, plutôt que sur le plan de la réalisation formelle de ces concepts dans les familles. Dans ce sens, la notion de *potentialité* est cruciale pour les paradigmes dérivationnels. Elle garantit un niveau de régularité et prévisibilité sur le plan conceptuel, dans le sens où les trous existants peuvent être remplis par de nouveaux mots (où plus précisément par des nouveaux lexèmes). Dans cette thèse nous rejoignons cette proposition. Par exemple, par rapport au paradigme en Table 2.6, sur le plan sémantique rien n’empêche d’avoir un lexème qui dénote un lieu dans une usine qui est aménagé au travail à la fraise ou au tour.

Pour récapituler, dans cette thèse nous cherchons à produire des paradigmes dérivationnels qui se composent de familles partielles alignées sur la base des relations sémantiques qu’elles contiennent. En outre, les familles organisées dans des paradigmes peuvent contenir des “trous” lorsque sur le plan conceptuel la case non remplie est bien motivée et disponible.

Malgré la caractérisation des paradigmes dérivationnels que nous venons de fournir, deux questions restent non résolues : **(a)** si les paradigmes sont des empilements des familles contenant les mêmes relations sémantiques, comment décidons-nous quelles sont les relations sémantiques pertinentes ? Pour le dire autrement, quand est-ce que deux familles peuvent être alignées sur

une base sémantique? **(b)** Si les paradigmes dérivationnels se composent des familles partielles, quel est le critère à utiliser pour délimiter ces familles? Ce problème dérive du problème de délimitation des familles partielles que nous avons introduit dans le Chapitre 1.

Cette thèse a comme objectif de fournir une méthode pour répondre à ces deux questions et de réaliser une délimitation des paradigmes dérivationnels. À ce propos, en Section 2.3, nous prenons en compte deux modèles paradigmatiques récents pour voir s'ils nous permettent d'apporter une réponse à ces questions : il s'agit des systèmes paradigmatiques de Bonami et Strnadová (2019) et du modèle *ParaDis* de Hathout et Namer (2022). Avant d'introduire ces deux modèles, nous présentons brièvement les apports principaux de l'approche paradigmatique de l'analyse des relations dérivationnelles.

2.2.4 Apport des paradigmes à l'analyse dérivationnelle

Les avantages d'une approche paradigmatique de la dérivation ont été discutés dans plusieurs travaux (Stump, 1991; Bauer, 1997; Štekauer, 2014; Hathout et Namer, 2019, 2022, entre autres). En premier lieu, les paradigmes permettent un traitement unifié de la flexion et de la dérivation. En deuxième lieu, les paradigmes en dérivation permettent de dépasser certaines limitations des analyses morphologiques basées sur des règles de construction des lexèmes (RCL) (Fradin, 2003), à savoir des règles binaires qui décrivent la relation entre une base et un dérivé. Contrairement aux RCLs, dans les modèles paradigmatiques de la dérivation les relations ne sont ni binaires ni orientées. Les paradigmes peuvent décrire n'importe quelle relation morphologique indépendamment de son orientation et de la distance morphologique entre deux lexèmes reliés dans une famille. Cela est utile pour décrire des relations entre lexèmes qui sont construits sur la même base et qui sont motivés sémantiquement. Par exemple, les paradigmes permettent un traitement des noms dérivés en *-ISME* et *-ISTE* en français, comme il est montré dans la Table 2.7.

La Table 2.7 présente des familles partielles où les noms construits en *-isme* et *-iste* sont dérivés d'une entité motivante X (qui peut être un nom commun ou un nom propre) et dénotent

respectivement une idéologie et un partisan de cette idéologie. La première ligne de cette table décrit le paradigme abstrait en précisant les propriétés formelles, catégorielles et sémantiques de chaque série. Les dérivés en *-iste* et *-isme* sont mutuellement motivés d’un point de vue sémantique : n’importe quelle idéologie a des partisans et ces partisans, à leur tour, se définissent par rapport à cette même idéologie. Cette relation reste présente aussi dans des familles qui ne contiennent pas d’entité motivante (cf. les deux dernières lignes de la table). Dans les familles de FASCISME ou de COMMUNISME, les noms en *-isme* et *-iste* n’ont pas de base commune mais sont toutefois motivés réciproquement. La réalisation des entités motivantes dans les familles de FASCISME et COMMUNISME ne se fait pas au niveau du lexique mais plutôt au niveau de la syntaxe. Les cases sont donc remplies par des périphrases comme “organisation militaire de la société et violente opposition à la démocratie” (i.e. valeurs du fascisme) et “mise en commun des moyens de production et abolition de la propriété privée” (i.e. valeurs du communisme).

X N [entité motivante]	Xisme N [idéologie]	Xiste N [partisan]
MARX	MARXISME	MARXISTE
TRUMP	TRUMPISME	TRUMPISTE
ESCLAVAGE	ESCLAVAGISME	ESCLAVAGISTE
-	FASCISME	FASCISTE
-	COMMUNISME	COMMUNISTE

TABLE 2.7 – Paradigme regroupant des familles contenant des noms dérivés en *-isme* and *-iste* mutuellement motivés (Hathout et Namer, 2019, p.7)

Les paradigmes dérivationnels permettent également de rendre compte des dérivés ayant une interprétation multiple. Par exemple, Strnadová (2014) utilise une approche paradigmatique pour la description des adjectifs relationnels (Adj_{rel}) en *-istique* en français, qui peuvent souvent donner lieu à de différentes interprétations. Les Adj_{rels} sont des adjectifs dénominaux dont le sens dépend de l’entité dénotée par leur nom base (Fradin, 2007; Strnadová, 2014; Namer, 2021).

Par exemple, JOURNALISTIQUE est construit sur JOURNALISTE, mais il peut être associé à d’autres noms de la même famille comme JOURNAL et JOURNALISME (Strnadová, 2014; Namer, 2021). Par conséquent, le sens de JOURNALISTIQUE n’est pas seulement lié sémantiquement à

JOURNALISTE, mais aussi à JOURNALISME et JOURNAL. Ce phénomène ne peut pas être décrit au moyen d'une RCL, qui se limiterait à décrire la relation entre JOURNALISTE (la base) et JOURNALISTIQUE (le dérivé).

Strnadová (2014) se sert du formalisme des *patrons cumulatifs* (*cumulative patterns* en anglais) proposé par Bochner (1993) pour décrire des relations multiples. Le modèle des patrons cumulatifs se base sur l'assomption que les relations dérivationnelles ne sont pas orientées. Un patron cumulatif est un ensemble d'éléments faisant partie de la même famille morphologique (flexionnelle ou dérivationnelle) qui sont interconnectés par des relations non orientées. Le patron cumulatif en Figure 2.3 est utilisé par Strnadová (2014) pour décrire l'interprétation multiple de JOURNALISTIQUE et sa motivation sémantique par rapport à JOURNAL et JOURNALISME. Ce patron cumulatif caractérise les lexèmes sur le plan formel, sur le plan catégoriel et sur le plan sémantique.

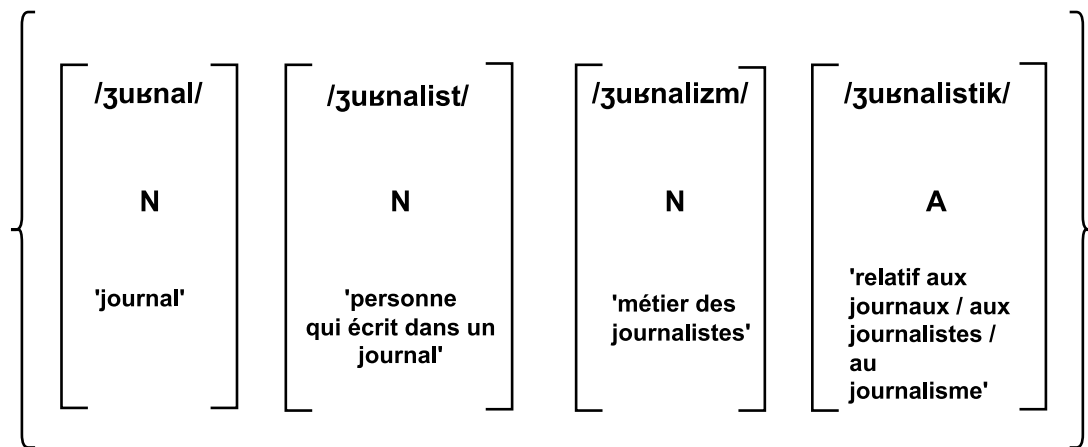


FIGURE 2.3 – Patron cumulatif proposé par (Strnadová, 2014, p.227) pour décrire le sens de JOURNALISTIQUE.

L'approche paradigmatique aux relations dérivationnelles permet également de décrire des cas de syncrétisme systématique au sein des familles dérivationnelles, comme par exemple le fait que le nom ITALIEN peut référer à l'Italie, à la langue italienne et à un habitant de l'Italie.

Pour récapituler, l'approche paradigmatique de la dérivation permet non seulement de représenter les régularités qui caractérisent le lexique construit mais aussi de dépasser certaines limites des RCLs dans la description de phénomènes comme ceux que nous venons d'exemplifier.

2.3 Exemples de modèles paradigmatiques

Nous avons conclu la Section 2.2.2 en soulignant deux questions qui restent irrésolues dans la définition des paradigmes dérivationnels : **(a)** si les relations sémantiques sont primaires pour la structuration des paradigmes en dérivation, comment peut-on établir que deux familles partielles contiennent des relations sémantiques alignables ? **(b)** si les paradigmes sont des empilements de familles, comment délimitons-nous les morceaux de familles que nous insérons dans le paradigme ? Pour pouvoir apporter une réponse à ces questions, nous nous intéressons aux modèles paradigmatiques existants pour voir si ces modèles peuvent nous aider. En particulier, dans cette section, nous nous concentrons sur deux modèles : les systèmes paradigmatiques de Bonami et Strnadová (2019) et le modèle *ParaDis* de Hathout et Namer (2022).

Le modèle proposé par Bonami et Strnadová (2019) est conçu pour pouvoir décrire à la fois des paradigmes en flexion et des paradigmes dérivationnels. Ce modèle se base sur la notion d'*alignement* des relations qui relie les membres des familles (flexionnelles et dérivationnelles) sur la base de leur contenu (morphosémantique ou morphosyntaxique) et considère les variations formelles comme étant secondaires pour la structuration des paradigmes. Cette approche est en ligne avec l'approche guidée par la sémantique proposée par Štekauer (2014) que nous avons introduite précédemment dans ce chapitre. Le modèle paradigmatique *ParaDis*, pour sa part, considère les régularités formelles, catégorielles et sémantiques comme étant indépendantes les unes des autres. Dans ce modèle, chaque niveau d'analyse linguistique (phonologique, catégoriel et sémantique) répond à ses propres contraintes. Cela est en ligne avec l'indépendance des réseaux conceptuels par rapport à leur réalisation dans le lexique, que nous avons introduite au Chapitre 1.

2.3.1 Les systèmes paradigmatiques de Bonami et Strnadová (2019)

Les systèmes paradigmatiques proposés par Bonami et Strnadová (2019) sont un modèle qui a été conçu pour représenter à la fois la flexion et la dérivation. Dans ce modèle, les familles (flexionnelles ou dérivationnelles) sont alignées sur la base des relations de contenu qui existent

entre leurs membres. La notion d'*alignement* joue donc un rôle central. Dans un paradigme, deux couples ordonnés de mots formes ou de lexèmes sont alignés s'ils sont reliés par la même relation de contenu. Les auteurs appellent cette relation *relation d'alignement* (*aligning relation* en anglais). La notion d'alignement peut donc être définie comme suit par Bonami et Strnadová (2019).

Définition 5**Alignement**

Deux couples ordonnés de mots (w_1, w_2) et (w_3, w_4) sont alignés s'il existe une relation de contenu R_c telle que $R_c(w_1, w_2)$ est égale à $R_c(w_3, w_4)$. Nous appelons R_c la relation d'alignement.

La Figure 2.4 représente un alignement de familles dérivationnelles partielles en français qui illustrent l'approche proposée par Bonami et Strnadová (2019) : si les relations de forme sont secondaires par rapport aux relations de sens pour la structuration des paradigmes en dérivation, alors on peut aligner les familles de LAVER, FORMER, LANCER et VOLER, qui sont structurées par le même réseau verbe-action-agent, indépendamment des différences au niveau formel que ces familles présentent.

Le paradigme présenté en Figure 2.4 est structuré par trois relations sémantiques qui sont définies comme suit :

- ' x est un nom qui dénote le même ensemble d'événements que le verbe y '
- ' x est un nom qui dénote un ensemble d'individus qui typiquement agissent en tant qu'agents dans les événements introduits par le verbe y '
- ' x est un nom qui dénote un ensemble d'individus qui typiquement agissent en tant qu'agents dans des événements introduits le nom y '

En revanche, Bonami et Strnadová (2019) mettent en évidence le fait que leurs systèmes paradigmatiques sont partiels et qu'il n'y a aucune contrainte ni sur la taille des familles qui peuvent être alignées dans un paradigme ni sur le nombre de familles qui peuvent être alignées dans le même système paradigmatique. Pour le dire autrement, un système paradigmatique peut

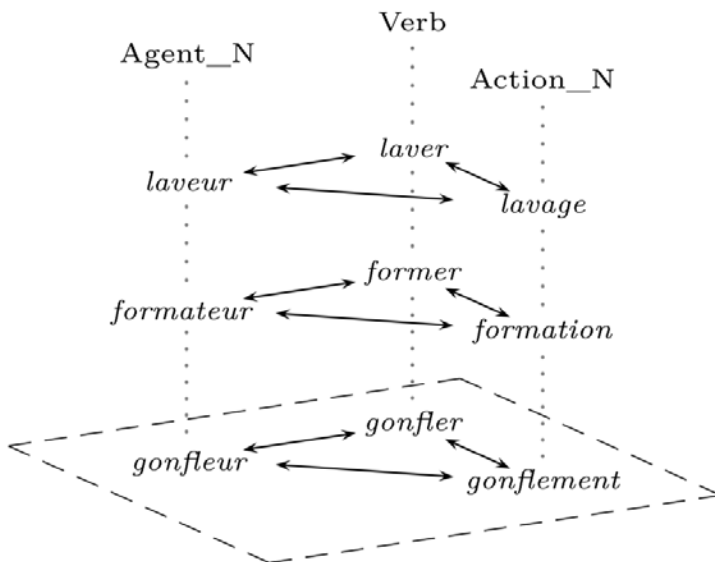


FIGURE 2.4 – Alignement de trois familles dérivationnelles partielles sur la base des relations sémantiques qui les structurent

contenir des familles qui ne sont pas exhaustives (i.e. qui ne sont pas complètes) et l'ensemble de familles qui sont alignées peut ne pas être exhaustif non plus.

Par conséquent, nous considérons que ce modèle n'apporte pas d'éléments qui peuvent nous aider à répondre à la question de la délimitation des paradigmes en dérivation ni précise quelles sont les conditions nécessaires à l'alignement entre deux relations. Par exemple, nous ne savons pas si au niveau sémantique le niveau de généralisation adopté pour former le système paradigmatique en Figure 2.4 est le plus pertinent pour décrire l'organisation du lexique construit dans l'esprit des locuteurs.

2.3.2 Le modèle paradigmatique ParaDis

Nous venons de voir dans que dans le modèle paradigmatique de Bonami et Strnadová (2019), les paradigmes dérivationnels sont structurés par des relations de contenu existantes entre lexèmes. Les relations sémantiques jouent un rôle important également dans le modèle *ParaDis* (Hathout et Namer, 2018; Namer et Hathout, 2020; Hathout et Namer, 2022), qui se

base sur l'idée que toutes les régularités pertinentes pour la morphologie sont paradigmatiques et doivent être considérées comme étant indépendantes l'une de l'autre (Hathout et Namer, 2022, p.167). *ParaDis* décrit familles et paradigmes sur trois niveaux de représentation distincts : un niveau formel, un niveau catégoriel et un niveau sémantique et en utilise un quatrième, le niveau morphologique, qui sert à connecter les trois niveaux que nous venons de citer. Malgré le fait que le niveau formel, catégoriel et sémantique soient interconnectés au moyen du niveau morphologique, ils répondent chacun à leurs propres contraintes (i.e. contraintes phonologiques, contraintes catégorielles et contraintes sémantiques). La Figure 2.5 représente la connexion des trois niveaux de représentation au moyen paradigme morphologique qui comprend les familles de LAVER et de SALER.

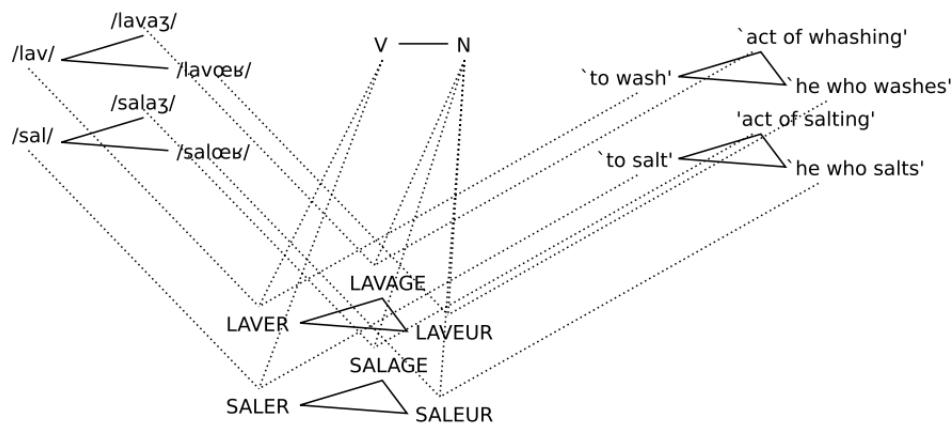


FIGURE 2.5 – Figure empruntée à (Hathout et Namer, 2022, p.172). Dans *ParaDis*, un paradigme morphologique est le résultat d'une correspondance d'un paradigme formel, un paradigme catégoriel et un paradigme sémantique. Dans le cas illustré dans cette figure, le paradigme phonologique se trouve sur la partie gauche de la figure, le paradigme catégoriel est placé au centre (V-N) et le paradigme sémantique est sur le côté droit. Le paradigme morphologique placé au centre en bas est le résultat de la convergence des trois paradigmes que nous venons de citer.

Dans *ParaDis* (Hathout et Namer, 2022), les sous-ensembles de lexèmes des familles qui s'alignent dans des paradigmes sont appelés *familles paradigmatiques* (*paradigmatic families* en anglais). Dans cette thèse, nous adoptons le même terme pour nous référer aux morceaux de familles dérivationnelles qui peuvent être alignés dans des paradigmes.

Définition 6**Famille paradigmatique**

Une famille paradigmatique est un sous-ensemble d'une famille dérivationnelle qui s'aligne avec d'autres familles paradigmatiques pour former un paradigme dérivationnel.

De la même manière que le modèle proposé par Bonami et Strnadová (2019), le modèle *ParaDis* ne précise pas comment délimiter les familles paradigmatiques. Toutefois, Hathout et Namer (2022) proposent de voir les paradigmes sémantiques pertinents pour la structuration des paradigmes dérivationnels comme des paquets de sens contenant des concepts interprédictibles. Ils considèrent que ces paquets de sens sont semblables à des scénarios cognitifs similaires aux cadres de la Sémantique des Cadres (*frame semantics* en anglais) (Fillmore *et al.*, 1976; Petruck, 1996; Baker, 2009; Ruppenhofer *et al.*, 2016). Dans cette thèse, nous parcourons cette piste pour proposer une méthode de délimitation des paradigmes dérivationnels sur base sémantique et nous nous inspirons aux scénarios définis par la Sémantique des Cadres. Ce cadre théorique et son implémentation dans la ressource lexicale *FrameNet* sont approfondis au Chapitre 3.

Pour résumer, les deux modèles paradigmatiques que nous venons de présenter, les *systèmes paradigmatiques* et *ParaDis*, ne posent pas de contraintes explicites sur la délimitation des familles paradigmatiques et donc ne peuvent pas être utilisés pour résoudre le problème de la délimitation des familles et des paradigmes. Il en va de même pour d'autres modèles paradigmatiques de la dérivation qui n'ont pas été détaillés dans ce chapitre, comme les patrons cumulatifs de Bochner (1993) ou le modèle proposé par Jackendoff et Audring (2019). À notre connaissance, la question de la délimitation des paradigmes en dérivation reste irrésolue dans la littérature.

Dans la prochaine section, nous cherchons des réponses possibles à cette problématique en prenant en compte des exemples de paradigmes dérivationnels qui ont été discutés dans des travaux récents.

2.4 Variété sémantique des paradigmes en dérivation

Comme nous l’avons vu précédemment, plusieurs travaux soutiennent l’idée que les relations sémantiques sont premières pour la structuration des paradigmes dérivationnels. En partant de ce principe, des travaux récents ont cherché à déterminer la nature de ces relations (Roché, 2017b; Fernández-Domínguez *et al.*, 2020; Fradin, 2021) et ont proposé une gamme de paradigmes dérivationnels présents dans des langues variées (Körtvélyessy *et al.*, 2020). Par exemple, Fradin (2020a) fait une première distinction entre paradigmes dérivationnels basés sur des relations de nature purement ontologique et un deuxième groupe de paradigmes basés sur des événements et leurs participants. Selon Fradin (2020a), les paradigmes structurés par des événements mettent en jeu des concepts tels que AGENT, PATIENT et INSTRUMENT qui peuvent difficilement être dissociés de l’événement dans lequel ils interviennent. Ces concepts correspondent en effet à des rôles sémantiques inscrits dans un événement (nous présentons plus en détail la notion de *rôle sémantique* dans le Chapitre 3). D’autre part, les paradigmes basés sur les propriétés ontologiques associées à un lexème de la famille ne peuvent pas être inscrits dans le cadre d’un événement. Dans d’autres cas encore, les paradigmes sont basés sur des relations qui ont leurs racines dans la société humaine, comme le paradigme basé sur les ethniques qui est discuté en Section 2.4.2. Dans cette section, nous présentons des exemples de paradigmes illustratifs des cas de figure que nous venons de représenter.

2.4.1 Paradigme basé sur des noms de fruits

Selon Fradin (2021), le paradigme basé sur les noms de fruit en français (Boyé et Schalchli, 2017; Roché, 2017b; Fradin, 2020a) est un exemple de paradigme complètement basé sur des relations de nature ontologique. Ce paradigme comprend des familles partielles comme celle en (17) : POMMIER dénote un arbre qui produit des pommes et POMMERAIE dénote un lieu planté de pommiers.

- (17) POMME.N, POMMIER.N, POMMERAIE.N

Selon Fradin (2021) le schéma abstrait du paradigme est structuré par des relations sémantiques reliant un fruit, la plante sur laquelle ce fruit pousse et le verger dans lequel la plante est plantée. Ces relations sont ancrées dans la réalité extralinguistique et découlent des propriétés ontologiques des fruits, à savoir le fait de pousser sur des plantes et que ces plantes poussent dans un terrain. Ces relations rentrent parmi les relations que (Koch, 1999) cite pour illustrer la notion de relation de contiguïté (voir Chapitre 1, Section 1.6.1) : un fruit, la plante qui le produit et le terrain sur lequel la plante pousse font partie du même scénario cognitif.

La Table 2.8 fournit un exemple de paradigme dérivationnel structuré par les relations sémantiques entre les trois concepts que nous venons d'introduire (i.e. FRUIT, PLANTE, TERRAIN). Selon Fradin (2021) ce paradigme contient des relations qui ne sont ni centrées sur l'humain (car il s'agit de propriétés naturelles des fruits) ni liées à un événement. Fradin (2021) souligne que le paradigme en Table 2.8 dépend principalement des propriétés ontologiques associées à l'entité dénotée par la base (i.e. le nom de fruit). Roché (2017b), pour sa part, parle de *élément initial* d'une famille pour se référer à l'élément de la famille qui n'est pas dérivé à partir d'autres éléments.

FRUIT	PLANTE	TERRAIN
POMME	POMMIER	POMMERAIE
AMANDE	AMANDIER	AMANDERAIE
CERISE	CERISIER	CERISAIE
PRUNE	PRUNIER	PRUNERAIE

TABLE 2.8 – Paradigme dérivationnel regroupant des familles qui contiennent un nom de fruit, le nom de la plante sur laquelle le fruit pousse et le nom du terrain où la plante est plantée.

Dans le paradigme en Table 2.8 ce sont principalement les propriétés de l'élément initial des familles qui assurent un certain niveau de prédictibilité sur le plan sémantique par rapport aux cases disponibles : si on a un nom qui dénote un fruit, alors il est possible que le nom de la plante sur laquelle ce fruit pousse et le nom du verger soient reliés morphologiquement au nom du fruit et fassent partie de la même famille. En français, cette réalisation est très régulière, comme le montrent les familles partielles présentées en Table 2.8.

Dans cette thèse nous allons voir d'autres exemples de paradigmes dont la structure se base principalement sur les propriétés sémantiques de l'élément initial de la famille. Par exemple, si on a dans une famille un nom de vêtement, alors dans la même famille nous pouvons trouver le nom qui dénote son fabricant et le nom qui dénote l'activité de fabrication (ex. BOTTE, BOTTIER, BOTTERIE). La notion d'élément initial de la famille joue donc un rôle important dans la plupart des paradigmes que nous allons présenter. Nous la définissons comme suit :

Définition 7**Élément initial d'une famille**

L'élément initial d'une famille dérivationnelle est un lexème qui n'est pas construit à partir d'autres membres de la famille.

Concernant le paradigme en Table 2.8, une question qui se pose est l'entité du réseau sémantique qui le structure. Le paradigme se structure autour de trois concepts, à savoir FRUIT, PLANTE et TERRAIN, mais la plupart des fruits sont produits dans le cadre d'activités humaines qui impliquent des AGRICULTEURS, des PRODUITS FERTILISANTS, etc. Est-ce que ces concepts font également partie du réseau sous-jacent au paradigme en Table 2.8? En outre, les agriculteurs et les activités de production fruiticole sont réalisés dans le lexique (spécialisé) au moyen de composés savants réguliers comme PRUNICULTURE, POMICULTURE, PRUNICULTEUR et POMICULTEUR. Est-ce que ces lexèmes peuvent être intégrés au paradigme en Table 2.8? Cette question est approfondie au Chapitre 5.

2.4.2 Paradigme basé sur les ethniques

Un autre paradigme dérivationnel qui a été discuté dans certains travaux récents est le paradigme des ethniques ou gentilés (Roché, 2008; Strnadová, 2014; Roché, 2017a; Schalchli et Boyé, 2018; Hathout *et al.*, 2021). Avec le terme *ethnique*, nous nous référons aux noms de peuples (ex. RUSSE), les noms d'humain construits sur un nom de pays ou de région (ex. ITALIEN, PIEMONTAIS) et les adjectifs relationnels correspondants (ex. ITALIEN.A, RUSSE.A, PIEMONTAIS.A). Schalchli et Boyé (2018), en reprenant la proposition de Roché (2008), considèrent

que la structure paradigmatique des ethniques repose sur un réseau comprenant trois catégories conceptuelles, à savoir une ZONE GÉOGRAPHIQUE (ville, région ou pays), une LANGUE et un HABITANT (i.e. l’ethnique). Ce paradigme est exemplifié en Table 2.9.

Contrairement à ce que nous avons vu en Section 2.4.2, il s’agit ici d’un paradigme qui ne dépend pas de propriétés naturelles d’une entité, mais plutôt de relations conceptuelles ancrées dans la société humaine. D’autre part, comme le paradigme basé sur les fruits, il s’agit d’un paradigme qui n’est pas structuré par un événement.

En outre, ce paradigme présente une autre particularité : il contient également les Adj_{rels} qui sont associés à chacun des trois sommets du réseau qui structurent le paradigme. L’approche paradigmatique de la description des adjectifs relationnels comme CHINOIS proposée par Schalchli et Boyé (2018) permet de rendre compte de leur plurivocité et du fait que ces Adj_{rels} peuvent être associés à l’habitant, à la zone géographique ou à la langue. Cette approche permet d’établir qu’il existe trois adjectifs relationnels distincts qui partagent la même forme, chacun associé à un sommet du réseau. Ils sont des éléments “sémantiquement annexes” aux sommets du réseau (ce point est approfondi au Chapitre 6). Cette analyse est similaire à celle réalisée par Strnadová (2014) pour JOURNALISTIQUE, qui peut s’associer à la fois à JOURNAL, à JOURNALISME et à JOURNALISTE (voir Section 2.2.4).

ETHNIQUE_N	ETHNIQUE_A	ZONE_N	ZONE_A	LANGUE_N	LANGUE_A
CHINOIS	CHINOIS	CHINE	CHINOIS	CHINOIS	CHINOIS
CORSE	CORSE	CORSE	CORSE	CORSE	CORSE
TADJIK	TADJIK	TADJIKISTAN	TADJIK	TADJIK	TADJIK

TABLE 2.9 – Exemple de paradigme dérivationnel basé sur les ethniques.

Le cas illustré dans la Table 2.9 montre que dans un même paradigme on peut avoir à la fois des relations entre participants d’un même scénario et des relations qui relient ces participants et l’adjectif qui en partage le contenu sémantique. Cette situation est schématisée en Figure 2.6 : nous représentons un réseau à trois sommets qui sont reliés par des relations de contiguïté. Chaque sommet, à son tour, est relié à son adjectif relationnel par une relation de nature orthogonale qui se définit indépendamment de la nature du réseau.

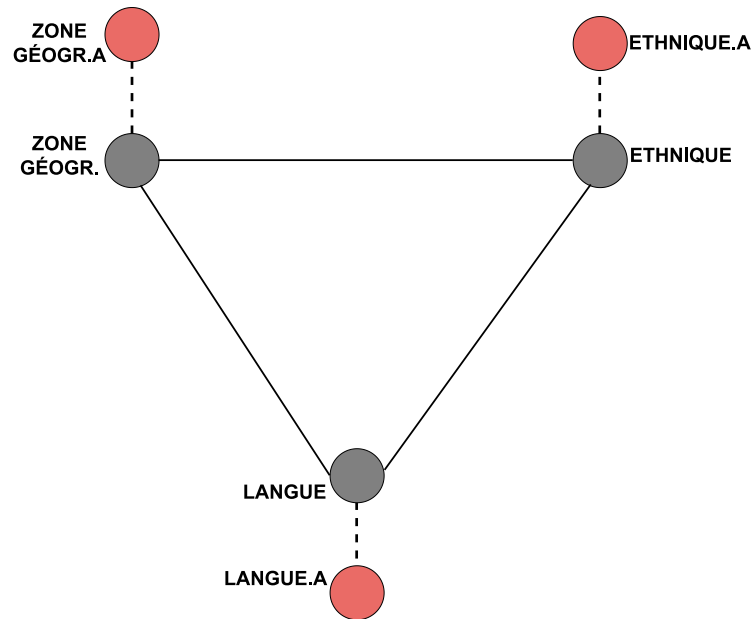


FIGURE 2.6 – Représentation sous forme de graphe du paradigme sémantique qui structure le paradigme dérivationnel en Table 2.9. Dans ce réseau on trouve des relations structurelles existantes entre les trois concepts précédemment cités et des relations de type orthogonal entre ces concepts et les adjectifs relationnels qui leur sont associés.

2.4.3 Réseau action

Les deux paradigmes que nous venons de présenter ne sont pas liés à des événements, mais sont basés soit sur les propriétés ontologiques des fruits (Table 2.8), soit sur des concepts liés à la société humaine (Table 2.9). Fradin (2021) montre qu’il existe également des paradigmes ancrés à un événement et qui comportent des concepts tels que AGENT, INSTRUMENT ou ACTION dont la caractérisation dépend de l’événement dans lequel ils sont impliqués. Il s’agit des rôles sémantiques (Gruber, 1965; Fillmore, 1967), qui ont été utilisés principalement dans la littérature pour décrire les relations entre un prédicat et ses arguments. Comme nous allons le voir plus en détail dans le Chapitre 3, l’attribution de ces rôles se fait généralement en contexte et il n’y a pas de consensus dans la littérature sur la définition de rôles comme AGENT ou INSTRUMENT. Par exemple, l’attribution du rôle d’AGENT implique des paramètres tels que la causalité, l’intentionnalité, le contrôle de l’événement, etc. Cela signifie que lorsqu’on identifie

un paradigme basé sur un événement, on doit tenir compte de tous les paramètres qui rentrent en jeu dans la définition des rôles sémantiques.

Un exemple de paradigme basé sur un événement est le *réseau action* proposé par (Roché, 2017b). Selon Roché, ce réseau se structure autour d’une action (i.e. le pivot sémantique du réseau) dont les participants correspondent aux autres sommets du réseau qui structure le paradigme. Le paradigme proposé en Table 2.10, par exemple, est structuré par un réseau qui contient une action (qui peut être réalisée dans le lexique sous forme verbale et nominale), un agent humain qui réalise cette action, un lieu dans lequel l’action se déroule et un instrument utilisé dans le processus (Fradin, 2021, p.159).

Les familles contenues dans le paradigme en Table 2.10 présentent des trous quand un concept donné qui est contenu dans le réseau action n’est pas réalisé morphologiquement (e.g. les lieux dédiés à l’action de TONDRE et de SOUDER). En outre, un concept du réseau peut être réalisé dans deux colonnes distinctes du paradigme : c’est le cas de l’action, qui peut être réalisée sous forme verbale et nominale. Le traitement de la relation entre verbes et noms d’action est décrit dans le Chapitre 6.

ACTION _ V	ACTION _ N	AGENT _ N	LIEU _ N	INSTRUMENT _ N
LAVÉR	LAVAGE	LAVEUR	LAVOIR	LAVEUSE
BINER	BINAGE	BINEUR	-	BINETTE
TONDRE	TONTE	TONDEUR	-	TONDEUSE
SOUDER	SOUDURE	SOUDEUR	SOUDERIE	SOUDEUSE

TABLE 2.10 – Paradigme dérivationnel structuré par le réseau action.

Un des problèmes mis en évidence par Fradin (2021) concernant l’applicabilité du réseau action pour délimiter les paradigmes en dérivation est que certains concepts (i.e. les participants de l’action) sont pertinents sur le plan sémantique uniquement pour des sous-ensembles d’actions. Par exemple, le sommet de l’instrument est pertinent uniquement pour des processus qui en impliquent un. En fonction du type d’événement dénoté par le verbe, le nombre de cellules disponibles dans le paradigme change. Par conséquent, Fradin (2021) constate que le réseau

action, conçu de cette manière, est trop général pour être efficace à la fois comme outils prédictifs par rapport aux cellules disponibles dans le paradigme et pour délimiter les paradigmes dérivationnels (Fradin, 2020a, p.70).

Pour améliorer l’efficacité du réseau action, Fradin (2021) propose de le voir comme un ensemble (ou *cluster*) de réseaux où les concepts qui sont disponibles dans le réseau sont déterminés par la structure argumentale du prédicat. La Table 2.11 montre des exemples de sous-réseaux : on distingue des actions intransitives (première ligne), des actions transitives (deuxième ligne), des actions transitives impliquant un instrument (troisième ligne) et des actions de création impliquant un agent, un instrument et un résultat (quatrième ligne). Dans d’autres actions encore, on peut avoir un LIEU dédié au procédé, comme LAVOIR et SOUDERIE.

VERBE	ACTION_N	AGENT_N	-	-	-
VERBE	ACTION_N	AGENT_N	PATIENT_N	-	-
VERBE	ACTION_N	AGENT_N	PATIENT_N	INSTRUMENT_N	-
VERBE	ACTION_N	AGENT_N	-	INSTRUMENT_N	RÉSULTAT
...

TABLE 2.11 – Des possibles sous-spécifications du réseau action, adaptée à partir de Fradin (2021, p.165). Chaque ligne représente un type distinct de réseau action. Chaque type de réseau action comporte une configuration différente au niveau des participants et, par conséquent, des cellules disponibles dans le paradigme.

La proposition de Fradin (2021) suggère que pour obtenir des paradigmes dérivationnels basés sur des événements qui ont un pouvoir prédictif au niveau sémantique, il faut adopter un niveau de granularité plus fin par rapport au réseau action (qui reste très général) et se baser sur des “classes d’actions”, comme par exemple les classes de verbes proposées par Levin (1993) ou les scénarios décrits par les cadres (Fillmore *et al.*, 1976). Dans cette thèse, nous parcourons cette piste et nous cherchons à déterminer le bon niveau de granularité pour obtenir des paradigmes dérivationnels sémantiquement délimités et stables au niveau des relations sémantiques qui les structurent.

2.4.4 Réseau activité

Dans son article sur les réseaux sémantiques qui structurent les familles dérivationnelles en français, Roché (2017b) propose également ce qu’il appelle *réseau activité*. Contrairement à ce que nous avons vu avec le réseau action, qui se structure autour d’un verbe, le *réseau activité* proposé par Roché (2017b) généralement se structure autour d’un nom d’activité massique qui est compatible avec la tournure *faire du/de la X* (18). Le réseau activité est motivé par le fait que les occupations humaines sont classables en divers types d’activité, alors que le réseau action met en avant un procès et ses participants (Fradin, 2021, p.166). Des exemples de familles structurées par un réseau de type activité sont fournis en (19).

- (18) a. *faire du jardinage*
 b. *faire de la boxe*
 c. *faire du journalisme*
- (19) a. JOURNAL.N, JOURNALISME.N, JOURNALISTE.N
 b. BIJOU.N, BIJOUTERIE.N, BIJOUTIER.N
 c. JARDIN.N, JARDINAGE.N, JARDINIER.N, JARDINER.V
 d. DANSE.N, DANSEUR.N, DANSER.V

Le réseau activité ne se base pas sur un événement ponctuel délimité dans le temps et dans l’espace, mais plutôt sur une série d’événements ou sur la répétition du même événement. Roché définit le nom d’activité comme « *une accumulation de procédés ou la répétition systématique du même procédé* » (Roché, 2017b; Fradin, 2021). Par exemple, POTERIE dénote une activité qui consiste à fabriquer régulièrement des pots. Lorsque les noms d’activité sont construits sur une base nominale, cette base nominale dénote généralement l’objet qui rend possible l’activité (ex. SKI) où un résultat de l’activité (JOURNAL, POT). Au sein du réseau activité, Roché (2017b) identifie les membres suivants :

- Une activité sous forme nominale (ex. SKI, JOURNALISME, POTERIE);

- Un agent “habituel” qui se consacre a cette activité et qui est nommé d’après celle-ci (ex. SKIEUR, JOURNALISTE) ;
- Un objet sur lequel l’activité est basée (ex. JOURNAL) ;
- L’activité sous forme verbale, lorsqu’elle est présente dans la famille (ex. JARDINER).

La Table 2.12 présente un exemple de paradigme structuré par le réseau activité que nous venons de présenter.

ACTIVITE _N	AGENT _N	OBJET	ACTIVITÉ _V
JARDINAGE	JARDINIER	JARDIN	JARDINER
BOXE	BOXEUR	-	-
VOILE	VOILIER	VOILE	-
JOURNALISME	JOURNALISTE	JOURNAL	-
BIJOUTERIE	BIJOUTIER	BIJOU	-

TABLE 2.12 – Exemple de paradigme dérivationnel structuré par le réseau activité proposé par Roché (2017b).

Cependant, Fradin (2021) souligne qu’un réseau activité comme celui que nous venons de présenter comporte les mêmes problèmes que le réseau action que nous avons présenté précédemment : il est trop général pour être vraiment opérationnel pour délimiter les paradigmes en dérivation. Fradin affirme que « *la valeur prédictive qui peut être associée au réseau action et au réseau activité est trop faible [...]. Il vaudrait la peine de voir si des réseaux plus stables ne peuvent pas être établis sur la base de domaines ou sous domaines d’activité* » (Fradin, 2020a, p.23).

Par exemple, la plupart des activités comportent des instruments et certaines activités sont nommées d’après l’instrument utilisé (ex. SKI, FLEURET), ou encore une partie d’activité se déroule dans un lieu dédié (ex. la BOTTERIE est un atelier de fabrication de bottes). En fonction du type d’activité considéré (ex. activités de manufacture, activités sportives), le réseau pourrait être donc plus ou moins étendu et comporter plus ou moins de cellules disponibles. Cette perspective est explorée dans les Chapitres 4 et 5, où nous proposons des paradigmes dérivationnels structurés par des domaines d’activités divers.

2.5 Bilan

Dans ce chapitre, nous avons introduit la notion de paradigme. Cette notion est principalement associée à la flexion, mais plusieurs travaux en proposent une adaptation à la dérivation pour décrire les régularités qui caractérisent le lexique construit (Štekauer, 2014; Bonami et Strnadová, 2019; Hathout et Namer, 2019). En outre, une approche paradigmatique de la dérivation permet de dépasser les limites de l’analyse basée sur des RCLs pour décrire la relation entre deux lexèmes qui n’ont pas de base commune (ex. FASCISME - FASCISTE) ou l’interprétation multiple d’adjectifs comme JOURNALISTIQUE.

Une des propriétés centrales pour les paradigmes est le fait qu’ils contiennent des éléments interprédictibles : cela signifie que dans un paradigme la présence d’un élément x prédit sur le plan morphosyntaxique (pour la flexion) ou morphosémantique (pour la dérivation) la présence d’un élément y et vice versa.

Le terme *paradigme* en morphologie dérivationnelle est utilisé avec des sens variés. Certains travaux l’utilisent pour se référer aux familles dérivationnelles (Bauer, 1997; Stump, 2019). En revanche, nous avons montré que les familles dérivationnelles complètes ne peuvent pas constituer des paradigmes car elles contiennent des éléments qui ne sont pas interprédictibles. En revanche, au sein des familles, on peut identifier des sous-ensembles de lexèmes qui fonctionnent en paquet et qui sont considérables comme étant interprédictibles.

Nous avons présenté une deuxième approche des paradigmes dérivationnels qui les considère comme étant des empilements de familles qui contiennent les mêmes relations. En particulier, nous nous sommes concentrés sur une approche qui considère que les paradigmes en dérivation sont principalement structurés par des relations sémantiques et que l’aspect formel est secondaire (Antoniova et Štekauer, 2016; Bonami et Strnadová, 2019). Dans cette thèse nous utilisons cette définition de paradigme dérivationnel.

Nous considérons que les relations sémantiques sont primaires pour la structuration des paradigmes et que les réseaux récurrents qui structurent les paradigmes sont indépendants de leur réalisation dans le lexique, comme le considèrent Hathout et Namer (2022) dans leur modèle

paradigmatique *ParaDis*.

D'autre part, nous avons montré que le problème de la délimitation des paradigmes en dérivation reste irrésolu dans la littérature. Ce problème dérive du problème de délimitation des familles dérivationnelles que nous avons introduit dans le Chapitre 1.

En outre, nous avons approfondi la question de la nature sémantique des paradigmes dérivationnels en nous appuyant sur des travaux récents (Roché, 2017b; Fradin, 2021). Nous avons introduit une première distinction, à savoir celle entre les paradigmes qui sont basés sur des événements (ex. réseau action) et des paradigmes qui ne le sont pas (ex. paradigme basé sur les ethniques). En outre, nous avons montré que le réseau action et le réseau activité proposés par Roché (2017b) sont trop généraux pour fonctionner comme des outils prédictifs et pour structurer des paradigmes dérivationnels stables sur le plan des relations sémantiques. Nous considérons que des paradigmes dérivationnels sémantiquement stables sont à chercher en prenant en compte des sous-domaines d'activité, comme il est suggéré par Fradin (2021). Cette question est approfondie dans les Chapitres 4 et 5.

Pour récapituler, nous considérons que les paradigmes dérivationnels sont structurés par des réseaux sémantiques, que ces réseaux sémantiques contiennent des concepts interprédictibles qui fonctionnent en paquets et que pour avoir des paradigmes sémantiquement stables nous devons prendre en compte des scénarios assez spécifiques (ex. domaines d'activité). Dans le Chapitre 3 nous introduisons le cadre théorique en sémantique que nous prenons comme inspiration pour apporter une réponse au problème de la délimitation des paradigmes : la Sémantique des Cadres (Fillmore *et al.*, 1976; Petruck, 1996; Baker *et al.*, 1998) ce cadre théorique se base justement sur des scénarios abstraits assez spécifiques qui regroupent des participants sémantiquement proches et interprédictibles.

Chapitre 3

Rôles thématiques et Sémantique des Cadres

Sommaire

3.1	Introduction	68
3.2	Les rôles sémantiques et leur usage en morphologie	69
3.3	Caractérisation sémantique des prédicats	74
3.4	Sémantique des Cadres et <i>FrameNet</i>	76
3.4.1	Modèle <i>FrameNet</i>	79
3.4.2	Rôles noyau	81
3.4.3	Caractérisation ontologique des rôles	82
3.4.4	Déclencheurs	83
3.4.5	Relations inter-cadre	85
3.5	Adaptabilité des cadres au découpage des familles	86
3.6	Bilan	87

BILAN DU CHAPITRE 1 : La notion de famille dérivationnelle joue un rôle central dans l'organisation du lexique construit. Cependant, la délimitation de cette notion reste un problème à résoudre dans la littérature (Hathout et Namer, 2019). Des travaux récents considèrent que les familles sont structurées par des réseaux conceptuels (Štekauer, 2014; Roché, 2017b). Au sein du système conceptuel, les concepts sont reliés par des relations de nature variée (taxinomiques,

de contiguïté) et s'organisent en réseau. L'interface entre lexique et système conceptuel a été discutée dans plusieurs travaux en sémantique lexicale : les relations lexicales reflètent en partie les réseaux contenus dans le système conceptuel (Cruse *et al.*, 1986; Murphy, 2003).

BILAN DU CHAPITRE 2 : L'organisation paradigmatique des relations dérivationnelles et du lexique construit a été discutée dans plusieurs travaux en morphologie dans les dernières décennies (Van Marle, 1985; Bauer, 1997; Hathout et Namer, 2019). Plusieurs auteurs considèrent que les paradigmes dérivationnels sont structurés par des relations sémantiques récurrentes et se composent de familles paradigmatiques contenant les mêmes relations (Bonami et Strnadová, 2019; Hathout et Namer, 2022). En revanche, la question de la délimitation des paradigmes en dérivation reste irrésolue et est liée à la problématique de la délimitation des familles.

3.1 Introduction

Dans ce chapitre, nous introduisons la Sémantique des Cadres (*frame semantics* en anglais) (Fillmore *et al.*, 1976; Petruck, 1996; Fillmore, 2006). Cette théorie a inspiré la méthodologie que nous proposons pour délimiter les paradigmes dérivationnels et que nous présentons dans le Chapitre 4. Nous prenons comme point de départ les hypothèses que nous avons introduites dans le Chapitre 2, à savoir le fait que les paradigmes dérivationnels sont en premier lieu structurés par des relations sémantiques récurrentes et que c'est par la sémantique que la délimitation des paradigmes en dérivation se fait.

En premier lieu, nous introduisons la notion de *rôle sémantique* (ex. AGENT, PATIENT, INSTRUMENT) (Gruber, 1965; Fillmore *et al.*, 1976)¹¹. Nous avons déjà pu voir ces rôles intervenir dans les exemples de paradigmes dérivationnels que nous avons fournis dans le Chapitre 2. En morphologie, les rôles sémantiques ont été utilisés dans plusieurs travaux pour décrire les propriétés sémantiques des noms construits, en particulier des noms déverbaux (Booij, 1986; Namer et Villoing, 2008; Rosenberg, 2008; Huyghe, 2021, entre autres). D'autre part, les listes

11. Nous notons les rôles sémantiques en majuscule, de la même manière que les concepts.

et les définitions des rôles sémantiques varient d'un auteur à l'autre et certains rôles sémantiques restent très génériques, ce qui les rend peu efficaces pour caractériser les relations entre les lexèmes des familles et les inscrire dans des paradigmes spécifiques comme les paradigmes que nous cherchons à obtenir.

Pour délimiter les familles dérivationnelles, nous nous inspirons d'un cadre théorique qui a été développé à partir de la notion de rôle sémantique : il s'agit de la Sémantique des Cadres (Fillmore *et al.*, 1976; Petruck, 1996; Baker *et al.*, 1998). La Sémantique des Cadres se sert de scénarios cognitifs dans lesquels des rôles sémantiques situationnels interviennent. Ce cadre théorique se place à un niveau de généralisation inférieur par rapport aux rôles sémantiques classiques. Les cadres permettent d'identifier des régularités sémantiques plus stables et des relations sémantiques qui sont interprédictibles en étant inscrites dans le même scénario cognitif.

Pour introduire les notions centrales de la Sémantique des Cadres, nous nous appuyons sur *FrameNet* (Baker *et al.*, 1998; Ruppenhofer *et al.*, 2016), la principale ressource lexicale qui implémente ce cadre théorique. Dans cette ressource, les cadres sont décrits par des sortes d'histoires abstraites qui se réalisent concrètement en discours, à savoir dans les corpus annotés sémantiquement.

Nous considérons que nous pouvons établir un ensemble de correspondances entre les éléments constitutifs de la Sémantique des Cadres et les réseaux sémantiques qui structurent les paradigmes dérivationnels et que nous cherchons à identifier.

3.2 Les rôles sémantiques et leur usage en morphologie

Dans le Chapitre 2, pour décrire les propriétés sémantiques de noms tels que *LAVEUR* et *JARDINIER*, nous avons utilisé la catégorie sémantique *AGENT*. Cette catégorie fait partie de ce qu'on appelle les *rôles sémantiques* (ou *rôles thématiques*). Les rôles sémantiques ont été initialement proposés par Gruber (1965) et Fillmore (1968) et ont été utilisés dans beaucoup de travaux pour décrire les relations sémantiques entre un prédicat et ses arguments syntaxiques (Jurafsky et Martin, 2014, Chapitre 24). Ces rôles fonctionnent comme des sortes d'étiquettes

que les prédicats attribuent à leurs arguments. Par exemple, si nous considérons l'énoncé en (20), nous pouvons identifier trois rôles : *Valentin* joue le rôle de l'AGENT, car il réalise une action de manière intentionnelle, *la vitre* joue le rôle du PATIENT, car elle est affectée par l'action et subit un changement d'état et le *marteau* joue le rôle de l'INSTRUMENT, car cet objet est utilisé comme intermédiaire pour réaliser l'action.

- (20) Valentin a cassé la vitre avec un marteau.
 AGENT PATIENT INSTRUMENT

Les rôles sémantiques permettent de faire des généralisations dans la description des relations entre un prédicat et ses arguments. Par exemple, si nous considérons un deuxième énoncé comme celui en (21), nous retrouvons les mêmes rôles que nous venons d'attribuer aux participants de la situation décrite en (20) : *le placard* joue le rôle de PATIENT, *Maxime* joue le rôle d'AGENT et le *tournevis* joue le rôle d'INSTRUMENT.

- (21) Le placard a été réparé par Maxime avec un tournevis.
 PATIENT AGENT INSTRUMENT

Les premiers travaux sur les rôles sémantiques avaient comme objectif la production d'une liste finie (et idéalement de petite taille) dans laquelle les rôles ont une définition qui est indépendante des prédicats qui les assignent. La Table 3.2 montre les principaux rôles sémantiques qui ont été discutés dans la littérature. Ces rôles sont accompagnés par une définition générale et un exemple d'attribution en contexte. Cette liste est purement illustrative : comme nous allons le voir, beaucoup de rôles (ex. AGENT et INSTRUMENT) ont fait objet de discussions dans la littérature et leurs définitions ne sont pas consensuelles.

Comme nous venons de l'annoncer, la définition et les critères d'attribution de beaucoup de rôles sémantiques ne sont pas consensuels. Tout le monde semble reconnaître l'existence d'un rôle d'AGENT, d'un rôle d'INSTRUMENT ou d'un rôle de THÈME, mais de nombreuses discussions ont été faites sur les critères à utiliser pour définir ces rôles. Par exemple, Fradin et Winterstein (2012) proposent de faire une distinction interne au rôle INSTRUMENT en distinguant l'INSTRUMENT du MOYEN : le premier est une entité qui existe séparément

Rôle sémantique	Définition	Exemple
AGENT	Initiateur de l'action, capable de volition	<i>Marc a cuisiné une soupe.</i>
BÉNÉFICIAIRE	Pour le bénéfice de qui l'action est accomplie	<i>Marie a acheté un cadeau pour sa soeur.</i>
CAUSE	Initiateur d'une action, non capable de volition	<i>Le feu a brûlé la cabane</i>
DESTINATION	Point d'arrivée d'un déplacement	<i>Nous avons pris le train pour Paris.</i>
EXPERIENCEUR	Perçoit l'action mais ne la contrôle pas	<i>Jean a une migraine.</i>
INSTRUMENT	Intermédiaire utilisé pour réaliser une action	<i>Max a cassé les pierres avec un marteau.</i>
PATIENT	Affecté par l'action, subit un changement d'état	<i>Julie a cassé la vitre</i>
RÉSULTAT	Produit final d'une action	<i>Claire a fabriqué un pot.</i>
SOURCE	Point de départ d'un déplacement	<i>Ce train arrive de Rome.</i>
STIMULUS	Produit un effet (psychologique) sur un expérienceur	<i>Le bruit a effrayé les passagers.</i>
THÈME	Entité déplacé ou localisée	<i>Alex a jeté le frisbee.</i>

TABLE 3.1 – Définitions générales et exemples d'attribution des rôles sémantiques plus utilisés.

avant et après l'événement dans lequel elle intervient (22a), contrairement au MOYEN qui subit un changement au cours du processus (22b).

- (22) a. Les chercheurs ont examiné cette protéine au microscope.
INSTRUMENT
- b. Claire a construit un mur avec des briques.
MOYEN

Pour prendre un autre exemple, la distinction entre AGENT et INSTRUMENT n'est pas toujours simple à déterminer. Fradin (2020a) souligne que certains noms comme SISMOGRAPHE dénotent des objets sophistiqués qui fonctionnent de manière autonome et qui peuvent réaliser des tâches qui ne sont pas accessibles à l'homme. Ces noms peuvent apparaître en position de sujet dans des énoncés comme celui en (23b). Booij (1986) propose d'étiqueter ces noms comme des PSEUDO-AGENTS.

- (23) a. Les chercheurs ont détecté le tremblement de terre à l'aide d'un sismographe.
- b. Le sismographe a détecté le tremblement de terre.

Concernant la morphologie dérivationnelle, les rôles sémantiques ont été réadaptés dans de nombreux travaux pour décrire la sémantique des noms déverbaux (Booij, 1986; Namer et Villoing, 2008; Rosenberg, 2008; Schnedecker et Aleksandrova, 2016; Huyghe, 2021) et plus en général pour décrire les propriétés sémantiques des noms dérivés. Huyghe et Tribout (2015,

p.99) soulignent que l'attribution des rôles sémantiques en morphologie vise une caractérisation sémantique intrinsèque des unités lexicales, déterminée hors discours. Par exemple, l'attribution de l'étiquette INSTRUMENT à un nom déverbal comme ARROSOIR signifie que ARROSOIR dénote une entité prototypiquement utilisée pour réaliser l'action dénotée par le verbe base ARROSER.

Une partie des études qui ont fait usage des rôles sémantiques a pris comme point de départ une opération affixale pour décrire toutes ses fonctions sémantiques : par exemple Schnedecker et Aleksandrova (2016) proposent une classification des noms en *-aire* et distinguent des noms de STATUT (ex. SIGNATAIRE), noms d'AUTEUR (ex. DÉMISSIONNAIRE) et des noms de BÉNÉFICIAIRE (ex. ABANDONNATAIRE). Pour leur part, Huyghe et Tribout (2015) montrent que les noms déverbaux en *-eur* peuvent rentrer dans la catégorie des noms d'AGENT (ex. MENTEUR) des noms d'INSTRUMENT (ex. ASPIRATEUR) ou dans les deux (ex. ENREGISTREUR). Parmi les noms d'AGENT ils distinguent des noms de STATUT (ex. TATOUEUR) des noms dénotant des AGENTS OCCASIONNELS (ex. AGRESSEUR) et des noms de DISPOSITION (ex. SÉDUCTEUR).

D'autres études se concentrent sur un rôle sémantique en décrivant toutes ses réalisations morphologiques : par exemple la ressource *VerbAgent* (Tribout *et al.*, 2012) contient des noms d'AGENT déverbaux formés au moyen de neuf suffixes différents (24). En outre, Huyghe et Wauquier (2020) soulignent que les noms d'AGENT peuvent être formés aussi à partir de nom (ex. BIJOUTIER) ou ne pas être dérivés (ex. PROPHÈTE).

- (24)
- a. *-eur* (DANSER → DANSEUR)
 - b. *-euse* (CHANTER → CHANTEUSE)
 - c. *-rice* (INSPECTER → INSPECTRICE)
 - d. *-eresse* (DÉFENDRE → DÉFENDERESSE)
 - e. *-aire* (CONTESTER → CONTESTATAIRE)
 - f. *-ant* (ATTAQUER → ATTAQUANT)
 - g. *-ante* (DIRIGER → DIRIGEANTE)

- h. *-ent* (ADHÉRER → ADHÉRENT)
- i. *-ente* (PRÉSIDER → PRÉSIDENTE)

L'étude de Salvadori et Huyghe (2023), pour sa part, porte sur la polyfonctionnalité affixale (i.e. la capacité des affixes de réaliser des fonctions sémantiques variées) en attribuant des rôles sémantiques aux noms déverbaux pour décrire leur relation avec le verbe base. Les rôles utilisés dans cette étude ont été adaptés à partir de *Verbnet* (Schuler, 2005), que nous présentons en Section 3.3, et par *LIRICS* (Petukhova *et al.*, 2008). La Table 3.2, présente des exemples d'attribution de rôles sémantiques à des noms déverbaux qui ont été utilisés dans la tâche de typage nominal réalisée dans le cadre de cette étude. Chaque rôle est accompagné par une définition proposée par les auteurs.

Rôle sémantique	Définition	Exemple
AGENT	Entité qui provoque intentionnellement un événement	FORGER → FORGERON
AMPLEUR	Valeur extensive liée à un événement	CONTENIR → CONTENANCE
BÉNÉFICIAIRE	Entité qui est avantagée par un événement ou un état	HÉRITER → HÉRITIER
CAUSE	Raison pour laquelle une éventualité se produit	AGGLUTINER → AGGLUTININE
CHEMIN	Trajectoire suivie lors d'un déplacement	DÉVIER → DÉVIATION
DESTINATION	Point d'arrivée d'un déplacement	CRACHER → CRACHOIR
EXPERIENCEUR	Perçoit l'action mais ne la contrôle pas	ADORER → ADORATEUR
LIEU	Entité qui sert de point de repère pour un événement	FUMER → FUMOIR
INSTRUMENT	Intermédiaire ou moyen utilisé pour réaliser une action	ARROSER → ARROSOIR
MANIÈRE	La façon dont une action est effectuée	PRONONCER → PRONONCIATION
PATIENT	Affecté par l'action, subit un changement d'état	MOURIR → MOURANT
RÉSULTAT	Produit final d'une action	ÉGRATIGNER → ÉGRATIGNURE
SOURCE	Point de départ d'un déplacement	GOUTTER → GOUTTIÈRE
PIVOT	Entité qui se voit attribuée une condition	POSSÉDER → POSSÉSSEUR
STIMULUS	Entité qui cause un état ou condition psychologique	EMMERDER → EMMERDEMENT
SUJET	Entité qui est le sujet d'une discussion ou activité cognitive	DÉVINER → DÉVINETTE
THÈME	Entité déplacé ou localisée	CHARGER → CHARGEMENT

TABLE 3.2 – Liste de rôles sémantiques utilisés par Salvadori et Huyghe (2023) pour le typage des noms déverbaux.

Comme on peut le voir, l'inventaire de rôles sémantiques en Table 3.2 est plus varié que celui en Table 3.1 et il est “calibré” pour les besoins de description morphosémantique des relations entre les noms déverbaux et leur base. En outre, dans leur étude les auteurs soulignent la nécessité de traiter de manière séparée les propriétés relationnelles des dérivés et leurs propriétés ontologiques. Par exemple, un nom déverbal qui appartient à la catégorie ontologique

ARTEFACT peut être relié par des relations différentes à sa base verbale : BÂTIMENT dénote un RÉSULTAT du prédicat dénoté par BÂTIR, alors que RASOIR dénote un INSTRUMENT qui intervient dans le prédicat dénoté par RASER et GARAGE dénote le LIEU aménagé à l'action de GARER (Salvadori et Huyghe, 2023, p.10). Nous adoptons également cette séparation entre propriétés ontologiques et propriétés relationnelles dans la description sémantique des paradigmes que nous proposons dans les Chapitres 4 et 5.

Tout en couvrant une large gamme de relations morphosémantiques, nous considérons que les étiquettes en Table 3.2 ne sont pas exploitables directement pour typer les dérivés des familles, car elles couvrent uniquement des relations entre verbes et noms déverbaux. Par exemple, elles ne seraient pas directement exploitables pour décrire les relations contenues dans les familles basées sur les noms de fruits que nous avons présentés dans le Chapitre 2 ou dans d'autres réseaux nominaux comme le triplet POT.N, POTIER.N, POTERIE.N (voir Chapitre 4) ou LOUP.N, LOUVE.N, LOUVETEAU.N (voir Chapitre 5). À notre connaissance, il n'y a pas d'inventaire général de rôles sémantiques calibrés sur les relations morphosémantiques du lexique construit français qui peut être utilisé pour classifier les relations contenues dans les familles dérivationnelles.

En outre, nous rappelons notre nécessité de déterminer les relations sémantiques (i.e. les rôles) qui coexistent souvent et qui sont intepredictibles. Sur cette base, une autre solution possible est de chercher dans la littérature des cadres théoriques qui proposent une caractérisation sémantique des prédicats dans lesquels interviennent plusieurs rôles en même temps. Les classes verbales de Levin (1993), que nous présentons en Section 3.3, et les cadres de Fillmore *et al.* (1976), que nous présentons dans la Section 3.4, s'inscrivent dans cette perspective.

3.3 Caractérisation sémantique des prédicats

En raison des problèmes de consensus sur une liste définie de rôles sémantiques dont nous avons parlé dans la section précédente (voir Table 3.1), plusieurs auteurs dans le domaine de la sémantique ont parcouru d'autres pistes pour la caractérisation des relations entre prédicats et arguments syntaxiques. Par exemple, Dowty (1991) propose une caractérisation des rôles

sémantiques encore plus générique en proposant la notion de *proto-rôles*, qui s’inspire à la notion de *prototype* de Rosch (1973). Il considère qu’il existe deux proto-rôles (i.e. PROTO-AGENT et PROTO-PATIENT) et qu’un argument syntaxique d’un verbe peut être un AGENT ou un PATIENT plus ou moins prototypique sur la base de ses propriétés (Tribout *et al.*, 2012, p.2520, pour un résumé des propriétés prototypiques des agents et des instruments proposées par Dowty). On passe donc d’une liste qui comprend une dizaine de rôles comme celle en Table 3.1 à deux rôles. Pour les raisons de spécificité que nous avons citées auparavant, la solution des proto-rôles ne peut pas être adaptée pour la description des relations morphosémantiques contenues dans les familles. En outre, les proto-rôles de Dowty ne fonctionnent qu’avec des prédicats dynamiques.

D’autres auteurs vont dans la direction opposée à celle prise par Dowty et proposent une caractérisation sémantique des prédicats qui comporte un inventaire plus large de rôles sémantiques. Par exemple, Levin (1993) considère que le comportement d’un verbe, particulièrement en ce qui concerne l’expression et l’interprétation de ses arguments, dépend en grande partie de son sens (Djemaa, 2017, p.16). Sur cette base, Levin groupe les verbes qui ont le même comportement syntaxique (ex. même nombre et type d’arguments) dans des classes. Un exemple de classe verbale est fourni en Figure 3.1 : cet exemple est extrait de *VerbNet* (Schuler, 2005), une ressource qui implémente les classes verbales de Levin pour la description du comportement syntaxique des verbes en anglais. Il existe également une version française de cette ressource (Danlos *et al.*, 2016), à savoir *VerbNet*, qui a été créé semi-automatiquement à partir de *VerbNet*.

La classe verbale DESTROY ‘détruire’ illustrée en Figure 3.1 contient des verbes comme ANNIHILATE ‘annihiler’, DEVASTATE ‘dévaster’ et RAVAGE ‘ravager’. Les verbes qui appartiennent à cette classe déclenchent trois constructions syntaxiques possibles, qui sont indiquées dans la section *frames* ‘cadres’ : *AGENT V PATIENT (avec INSTRUMENT)*, *AGENT V PATIENT*, *INSTRUMENT V PATIENT*. De manière générale, les rôles sémantiques utilisés par *VerbNet* correspondent aux rôles sémantiques classiques tout en étant associés aux situations correspondant aux classes verbales.

Go To COMMENTS		destroy-44 <small>Members: 31, Frames: 3</small>	
MEMBERS			
ANNIHILATE (FN 1; WN 1)	DEVASTATE (FN 1; WN 1; G 1)	MUTILATE (FN 1; WN 1, 2, 3)	
BLIGHT (WN 1)	DISFIGURE (WN 1)	OBLITERATE (FN 1; WN 4; G 1)	
BLITZ (WN 1)	EFFACE (WN 3; G 3)	RACK (WN 8; G 1)	
DAMAGE (FN 1; WN 1; G 1)	EXTERMINATE (FN 1; WN 1; G 1)	RAVAGE (WN 2)	
DECIMATE (WN 2)	EXTIRPATE (WN 1)	RAZE (FN 1; WN 1)	
DEMOLISH (FN 1; WN 1; G 1)	LEVEL (FN 1)	RUIN (FN 1; WN 1; G 1)	
DESECRATE (WN 1; 2)	LOUSE_UP	SHATTER (FN 1; WN 1, 3; G 1)	
DESTROY (FN 1; WN 1, 2; G 1)	MAIM (WN 1)	SHIPWRECK (WN 1; G 1)	
ROLES			
<ul style="list-style-type: none"> • AGENT [+INT_CONTROL] • PATIENT [+?CONCRETE] • INSTRUMENT [+CONCRETE] 			
FRAMES			
NP V NP			
EXAMPLE	"The Romans destroyed the city."		
SYNTAX	AGENT V PATIENT		
SEMANTICS	CAUSE(AGENT, E) DESTROYED(RESLT(E), PATIENT)		
NP V NP PP INSTRUMENT			
EXAMPLE	"The builders destroyed the warehouse with explosives."		
SYNTAX	AGENT V PATIENT (WITH) INSTRUMENT		
SEMANTICS	CAUSE(AGENT, E) USE(DURING(E), AGENT, INSTRUMENT) DESTROYED(RESLT(E), PATIENT)		
NP INSTRUMENT V NP			
EXAMPLE	"The explosives destroyed the warehouse."		
SYNTAX	INSTRUMENT V PATIENT		
SEMANTICS	CAUSE(?AGENT, E) USE(DURING(E), ?AGENT, INSTRUMENT) DESTROYED(RESLT(E), PATIENT)		

FIGURE 3.1 – Classe verbale DESTROY présentée dans *Verbnet*(Schuler, 2005). Cette classe verbale comporte une situation impliquant un AGENT, un PATIENT et éventuellement un INSTRUMENT.

3.4 Sémantique des Cadres et *FrameNet*

Pour sa part, la proposition de Fillmore *et al.* (1976) se base sur la notion de *scénario*. La Sémantique des Cadres, initiée par le travail de Charles Fillmore (Fillmore, 1977, 1985; Fillmore et Atkins, 1992), est une théorie cognitive qui voit le sens linguistique comme une manifestation de la structure conceptuelle (Evans, 2006, p.156). Fillmore inscrit les rôles sémantiques dans des situations stéréotypées qu’il appelle *cadres* en s’inspirant à la notion de *cadre* utilisée par Minsky (1974) dans le domaine de l’intelligence artificielle. Nous précisons que la notion de cadre a été utilisée dans des domaines variés avec des sens et des fonctions divers : par exemple, Barsalou (2012) et d’autres auteurs qui ont repris ses travaux (Löbner, 2013; Petersen et Gamerschlag, 2014) utilisent le terme *cadre* pour se référer à tout “morceau de connaissance” (*chunk of knowledge* en anglais (Löbner, 2021, p.262)) qui fait partie de la cognition humaine sans se baser sur la notion de *scénario*. Dans notre discussion sur la notion de *cadre*, nous nous concentrons sur

entre sémantique et syntaxe au moyen de situations cognitives (i.e. les cadres) et inscrit dans ces situations des rôles sémantiques spécifiques qui sont interprédictibles.

Nous considérons que la notion de *cadre*, telle qu'elle est présentée par Fillmore, peut être adaptée à notre objectif de délimitation des paradigmes dérivationnels. En premier lieu, les cadres décrivent des scénarios qui sont indépendants de leur réalisation linguistique (en discours ou dans le lexique), comme les réseaux sémantiques qui structurent les paradigmes. En outre, dans les cadres les participants sont interprédictibles, comme les éléments des paradigmes que nous cherchons à produire. Par exemple, si dans un cadre nous avons un ACHETEUR, il va très probablement y avoir aussi un VENDEUR qui est dans le même scénario. Sur cette base, nous pouvons envisager de baser la délimitation des familles sur la base de situations sous-jacentes similaires aux cadres de Fillmore. Par exemple, si nous reprenons la famille en Figure 3.2 que nous avons introduite au Chapitre 1, nous pouvons considérer que les relations qu'elles contiennent s'inscrivent dans deux scénarios distincts qui correspondent à deux cadres distincts : POT, POTIER et POTERIE s'inscrivent dans un scénario de fabrication de pots, où on s'attend à avoir aussi des MATÉRIAUX des INSTRUMENTS DE FABRICATION (ex. tour) et potentiellement un LIEU DE FABRICATION (ex. un atelier). Pour leur part, les relations qui relient POT à POTIER et REMPOTEUR s'inscrivent dans un autre cadre qui décrit l'usage du pot en tant que RÉCIPIENT : dans ce même cadre on s'attend avoir une ENTITÉ DÉPLACÉE (généralement une plante) et un AGENT DU DÉPLACEMENT qui réalise l'action de REMPOTER.

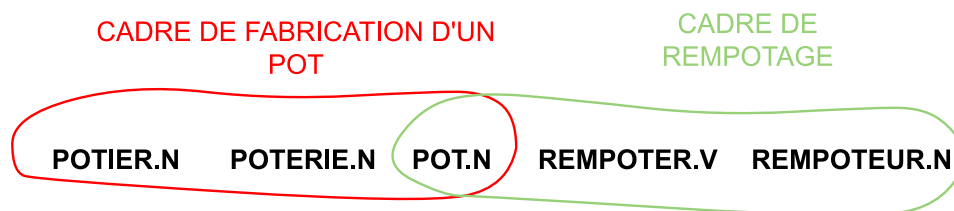


FIGURE 3.2 – Découpage de la famille de POT basé sur deux cadres distincts.

Les principes fondamentaux de la Sémantique des Cadres que nous venons d'évoquer ont été mis en œuvre dans plusieurs programmes de recherche pour la création de ressources lexicales à grande échelle pour l'annotation sémantique des textes. La ressource principale qui implémente

les principes de la sémantique des cadres est *FrameNet* (Lowe, 1997; Baker *et al.*, 1998; Boas, 2005; Ruppenhofer *et al.*, 2016; Baker, 2017), une ressource pour l’annotation de corpus anglais basée sur les cadres. Au cours des dernières décennies, plusieurs ressources conçues sur le modèle *FrameNet* ont été construites pour l’annotation sémantique de corpus dans d’autres langues (Subirats et Sato, 2003; Ohara *et al.*, 2004; Candito *et al.*, 2014; Borin *et al.*, 2021, parmi d’autres) et plusieurs travaux et conférences ont été dédiés au projet d’un *FrameNet* multilingue (Gilardi et Baker, 2018; Torrent *et al.*, 2018, 2020). En dernier lieu, d’autres ressources basées sur le modèle *FrameNet* ont été également été créés pour l’annotation de corpus de domaine spécialisé (Dolbey *et al.*, 2006; Schmidt, 2009; Venturi *et al.*, 2009).

Pour illustrer comment fonctionnent les cadres, leur structure interne et quels aspects peuvent être adaptés à notre objectif de délimitation des paradigmes dérivationnels, nous nous appuyons sur des exemples extraits par *FrameNet* (Baker *et al.*, 1998; Ruppenhofer *et al.*, 2016).

3.4.1 Modèle *FrameNet*

Comme nous venons de le voir, la théorie de la Sémantique des Cadres se base sur deux objets principaux : les cadres et les rôles à l’intérieur des cadres. Les ressources basées sur le modèle *FrameNet* disposent d’une grande quantité de cadres créés par des experts et qui sont utilisés principalement pour l’annotation sémantique de corpus. Elles sont généralement structurées en trois volets : **(a)** les cadres forment un ensemble structuré de situations prototypiques; **(b)** chaque cadre est associé à un ensemble d’unités lexicales qui évoquent ces cadres, que l’on appelle les *déclencheurs*; **(c)** chaque cadre est utilisé pour annoter des énoncés contenus dans des corpus. Pour donner une idée de la portée d’une ressource comme *FrameNet*, la version 1.5 comporte 11829 déclencheurs distincts pouvant évoquer 1019 cadres.

Nous considérons un exemple de cadre pris par *FrameNet* en Figure 3.3. En premier lieu, un cadre est décrit par une définition phrastique, appelé *description du cadre* (*frame description* en anglais). La description du cadre est une sorte d’histoire prototypique abstraite qui présente en même temps les participants principaux qui interviennent dans la situation ou dans l’objet décrit

par le cadre. Aucune contrainte formelle n'est appliquée sur la définition d'un cadre : il s'agit d'une définition en prose. Par exemple, la définition du cadre TUER est "Un MEURTRIER ou une CAUSE cause la mort d'une VICTIME". En outre, les éléments du cadre ou rôles du cadre sont également associés à une définition partielle qui peut impliquer d'autres rôles. Par exemple, le MEURTRIER est défini comme "la personne ou l'entité qui cause la mort de la victime". Nous précisons que l'usage de rôles "spécifiques" aux cadres dans *FrameNet* n'est pas systématique : par exemple, dans le cadre TUER en Figure 3.3 on trouve parmi les rôles du cadre le rôle INSTRUMENT, alors qu'un rôle ARME n'est pas présent dans ce même cadre.

Killing	
Definition:	
A Killer or Cause causes the death of the Victim . John DROWNED Martha .	
FES:	
Core:	
Cause □	An inanimate entity or process that causes the death of the Victim . The rockslide KILLED nearly half of the climbers.
Excludes: Killer	
Instrument [Instr]	The device used by the Killer to bring about the death of the Victim . It's difficult to SUICIDE with only a pocketknife.
Semantic Type: Physical_entity	
Excludes: Cause	
Killer [Kill]	The person or sentient entity that causes the death of the Victim .
Excludes: Cause	
Means □	The method or action that the Killer or Cause performs resulting in the death of the Victim . The flood EXTERMINATED the rats by cutting off access to food.
Semantic Type: State_of_affairs	
Excludes: Cause	
Victim □	The living entity that dies as a result of the killing.
Semantic Type: Sentient	

FIGURE 3.3 – Définition et rôles principaux du cadre TUER présent dans *FrameNet*.

Dans *FrameNet*, la définition du cadre et les définitions associées aux rôles du cadre sont accompagnées par des exemples annotés qui montrent la réalisation syntaxique des participants du cadre. Les exemples de corpus annotés sont donc des instanciations concrètes de la situation

décrite par le cadre. Pour illustrer la réalisation syntaxique des rôles du cadre, l'interface de *FrameNet* utilise des couleurs, comme nous le montrons en Figure 3.3 et dans l'exemple en (26).

- (26) John a noyé Martha.
 MEURTRIER VICTIME

3.4.2 Rôles noyau

FrameNet distingue les rôles en fonction de leur centralité dans la situation cognitive (*coreness*). En premier lieu, on distingue les *rôles noyau*, à savoir des rôles qui représentent un composant conceptuellement nécessaire pour le cadre et qui le rend unique et différent des autres (Djemaa, 2017, p.35). Comme on peut le voir dans la Figure 3.3, pour le cadre TUER les éléments qui sont considérés comme étant conceptuellement nécessaires pour la réalisation de la situation sont le MEURTRIER, la VICTIME, l'INSTRUMENT et le MOYEN. Nous précisons que ces rôles sont nécessaires uniquement sur le plan conceptuel et non sur le plan syntaxique. En outre, ce cadre contient un cinquième rôle noyau, la CAUSE, qui est reliée par une relation d'exclusion au MEURTRIER : cela parce que la situation décrite par le cadre TUER peut aussi être instanciée par des situations concrètes dans lesquelles une cause naturelle cause la mort d'une victime, plutôt qu'une personne qui tue la victime. Une liste détaillée des conditions de nuclearité pour un rôle sémantique est fournie par Ruppenhofer *et al.* (2016).

En outre, *FrameNet* distingue deux autres types de rôles : *périphériques* et *extra-thématiques*, même si la distinction entre les deux n'est pas toujours claire (Djemaa, 2017, p.35). L'interface graphique utilisée par *FrameNet* sépare les rôles sémantiques selon leur statut de rôle noyau ou non, sans indiquer quels sont les rôles périphériques et quels sont les rôles extra-thématiques. Des exemples de rôles sémantiques non-noyau qui sont inscrits dans le cadre TUER sont proposés en Figure 3.4 : on trouve des rôles comme LIEU ou TEMPS qui ne sont pas spécifiques à un cadre qui décrit un meurtre, mais sont pertinents pour n'importe quel cadre qui décrit des événements délimités dans le temps et dans l'espace.

Comme nous allons le voir dans le Chapitre 6, la distinction entre rôles noyau vs non-noyau (ou entre rôles centraux et rôles périphériques) peut être adaptée en morphologie pour

<p>Place [Loc] Semantic Type: Locative_relation</p>	<p>The location where the death took place. Apparently there were two MURDERS in Atherton in one week.</p>
<p>Purpose [Purp] Semantic Type: State_of_affairs</p>	<p>The state of affairs that the Killer is trying to bring about by killing. It's simple: he KILLED her to get the inheritance.</p>
<p>Result [R]</p>	<p>The state of the Victim that results from being killed. (Normally this is somewhat redundant.) Two men ... SMOTHERED her to death.</p>
<p>Time [T] Semantic Type: Time</p>	<p>The point or period of time within which the Victim is killed. There is no evidence that the royal family was KILLED during the destruction at this site.</p>

FIGURE 3.4 – Rôles non-noyau (i.e. *non core frame elements*) qui apparaissent dans le cadre TUER dans *FrameNet*.

distinguer des relations qui structurent un réseau conceptuel donné pertinent pour délimiter un paradigme et d'autres relations que l'on trouve presque systématiquement dans les familles, mais qui répondent à des besoins morphologiques ou syntaxiques plutôt qu'à un besoin conceptuel. Deux exemples de relations de ce deuxième type sont les relations entre verbes et noms dénotant le même prédicat (ex. LAVER - LAVAGE) et les relations entre noms et adjectifs relationnels (ex. ÉLECTION - ÉLECTORAL).

3.4.3 Caractérisation ontologique des rôles

FrameNet se sert également d'étiquettes ontologiques qui peuvent être attribuées aux rôles du cadre. Les étiquettes attribuées aux rôles du cadre ajoutent des restrictions de type ontologique sur les remplisseurs du rôle. Pour le cadre TUER en Figure 3.3, par exemple, l'étiquette SENTIENT est attribuéé au rôle de la VICTIME, ce qui veut dire que les constituants qui réalisent ce rôle en discours doivent référer à une entité douée de perception. De son côté, le rôle INSTRUMENT doit être rempli par des éléments syntaxiques qui font référence à une ENTITÉ PHYSIQUE. Les étiquettes attribuées aux rôles sont extraites par la hiérarchie de types indiquée par Lönneker-Rodman et Baker (2009). Toutefois, comme le souligne Djemaa (2017), le typage ontologique des rôles n'est pas systématique ni majoritaire dans *FrameNet*. En outre,

la version complète de cette ontologie ne semble pas être disponible sur le site de la ressource ni illustrée dans un des nombreux articles qui présentent *FrameNet*.

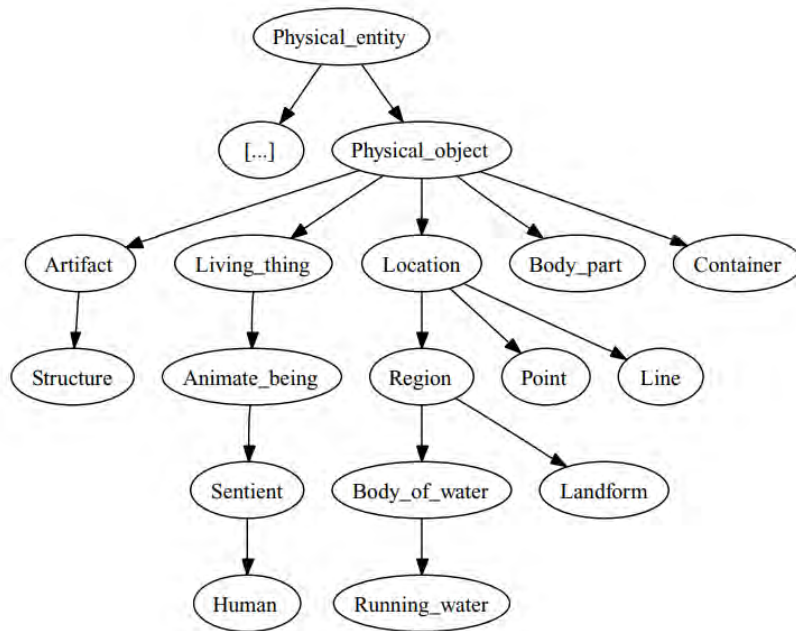


FIGURE 3.5 – Extrait des types ontologiques utilisés dans *FrameNet* pour appliquer des restrictions sur les rôles des cadres (Lönneker-Rodman et Baker, 2009, p.422).

Tout en n'étant pas systématique, *FrameNet* décrit la nature ontologique des participants à l'intérieur des relations. Cela est compatible avec notre nécessité de traiter de manière séparée les propriétés sémantiques et les propriétés relationnelles des lexèmes dans les familles. En revanche, l'inventaire des étiquettes ontologiques utilisées par *FrameNet* ne semble pas être disponible et ne peut pas être adapté à notre tâche.

3.4.4 Déclencheurs

Le *déclencheur* (*lexical unit* dans *FrameNet*) est défini comme étant l'élément syntaxique qui évoque le cadre dans l'esprit du lecteur/auditeur. Les ressources construites sur le modèle *FrameNet* présentent pour chaque cadre une liste de déclencheurs. La Figure 3.6 montre un

échantillon des déclencheurs du cadre TUER, qui sont listés dans la rubrique *lexical units*. Comme on peut le voir, un cadre peut être évoqué par des déclencheurs appartenant à des catégories syntaxiques variées. Pour leur part, la rubrique *FE Core set(s)* indique quels sont les rôles noyau pour le cadre TUER et la rubrique *frame to frame relations* décrit quelles sont les relations qui relient le cadre TUER à d'autres cadres dans *FrameNet* (ex. le cadre TUER est causatif du cadre MORT). Les relations inter-cadre sont approfondies en Section 3.4.5.

FE Core set(s):

{Instrument, Killer, Means}

Frame-frame Relations:

Inherits from: [Transitive action](#)

Is Inherited by: [Execution](#)

Perspective on:

Is Perspectivized in:

Uses:

Is Used by: [Surviving](#)

Subframe of:

Has Subframe(s):

Precedes:

Is Preceded by:

Is Inchoative of:

Is Causative of: [Death](#)

See also:

Lexical Units:

*annihilate.v, annihilation.n, asphyxiate.v, assassin.n, assassinate.v, assassination.n, behead.v, beheading.n, bl
decapitation.n, destroy.v, dispatch.v, do.in.v, drown.v, eliminate.v, euthanasia.n, euthanize.v, exterminate.v, ex
infanticide.n, kill.v, killer.n, killing.n, lethal.a, liquidate.v, liquidation.n, liquidator.n, lynch.v, massacre.n, mas
slaughter.n, slaughter.v, slaughterer.n, slay.v, slayer.n, slaying.n, smother.v, smothering.n, starve.v, suffocate.v,*

FIGURE 3.6 – Rôles noyaux, relations inter-cadres et déclencheurs du cadre TUER. La présentation d'un cadre dans *FrameNet* présente aussi les relations qui relient un cadre aux autres cadres (voir Section 3.4.5.)

En revenant aux déclencheurs, Ruppenhofer *et al.* (2016) souligne le fait qu'un mot polysémique peut évoquer des cadres distincts dans *FrameNet*. Il fait l'exemple du verbe anglais *to bake* 'cuire', qui évoque trois cadres distincts : APPLY HEAT 'appliquer de la chaleur'(27a), COOKING CREATION (27b) 'création culinaire' et ABSORB HEAT (27c) 'absorber la chaleur'.

- (27) a. *Michelle baked the potatoes for 45 minutes.*

'Michelle a fait cuire les pommes de terre pendant 45 minutes'

- b. *Michelle baked her mother a cake for her birthday.*
 ‘Michelle a préparé un gâteau pour l’anniversaire de sa mère.’
- c. *The potatoes have to bake for more than 50 minutes*
 ‘Les pommes de terre doivent cuire pendant plus de 50 minutes.’

3.4.5 Relations inter-cadre

Comme nous l’avons annoncé précédemment, *FrameNet* est un ensemble structuré de cadres : cette ressource se structure comme un réseau dans lequel les cadres sont connectés au moyen de relations. La fonction principale des relations inter-cadre est de rendre compte des relations entre les situations cognitives que les cadres décrivent. Les relations qui relient les cadres dans *FrameNet* rendent compte de phénomènes comme l’héritage, la subsumption, la causalité ou la précédence temporelle entre les situations ou les objets décrits par les cadres. La liste des types de relations inter-cadre est fournie par Ruppenhofer *et al.* (2016, p.75). Par exemple, dans *FrameNet* le cadre PROCÈS est un sous-cadre du cadre PROCÉDURE PÉNALE : un procès est une partie de la procédure pénale, qui comprend aussi l’état d’arrestation, la décision du tribunal, etc. La relation plus fréquemment utilisée dans *FrameNet* est la relation d’héritage qui connecte un cadre superordonné et un cadre subordonné, qui en hérite des propriétés ou des participants. Par exemple, dans *FrameNet* le cadre VENGEANCE est subordonné par rapport au cadre RÉCOMPENSES ET PUNITIONS.

Nous considérons que les réseaux sémantiques qui structurent les paradigmes en dérivation peuvent également être reliés par des relations inter-réseau. Nous illustrons cette considération en prenant en compte les deux sous-ensembles en (28).

- (28) a. LOUP.N, LOUVE.N, LOUVETEAU.N
 b. LOUP.N, LOUVETIER.N

Nous considérons que les relations sémantiques qui structurent le sous-ensemble en (28a) sont internes à l’espèce animale des loups : ces relations s’établissent entre exemplaires mâle, femelle et jeune de la même espèce. D’autre part, dans le couple en (28b) LOUP est en relation avec

LOUVETIER en tant qu'espèce animale et non en tant qu'exemplaire mâle, femelle ou petit. Nous considérons que ces deux sous-ensembles sont inscrits dans deux réseaux conceptuels distincts, mais que ces deux réseaux peuvent être reliés au moyen d'une relation de type EST-UN comme nous le montrons en Figure 3.7. Un loup, une louve et un louveteau appartiennent tous à l'espèce animale des loups (i.e. *canis lupus*) et l'espèce est à son tour reliée au chasseur spécialisé dans l'élimination de cet animal.

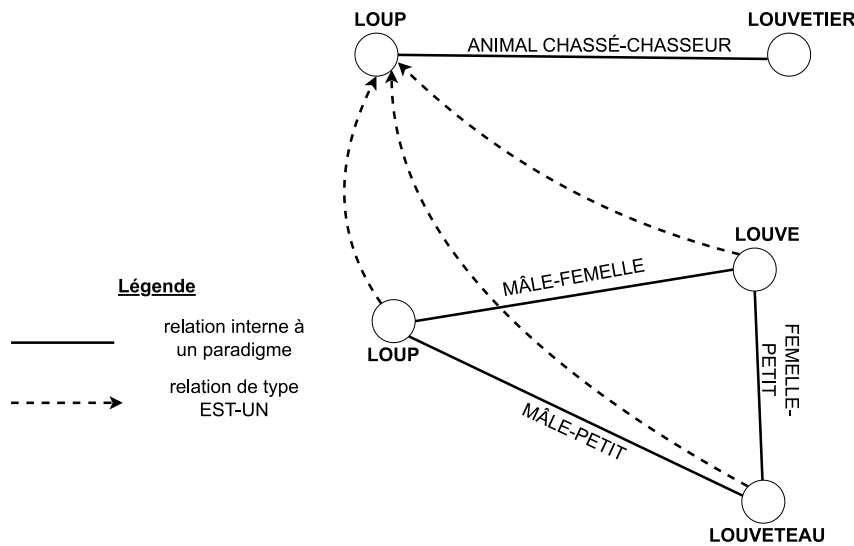


FIGURE 3.7 – Exemple de relations reliant deux réseaux conceptuels distincts.

La possibilité de relier les réseaux sémantiques qui structurent les paradigmes dérivationnels au moyen de relations est approfondie dans le Chapitre 5.

3.5 Adaptabilité des cadres au découpage des familles

Dans la section précédente, nous avons présenté les propriétés principales des cadres dans la conception de Fillmore en nous appuyant sur leur implémentation dans *FrameNet*. Nous considérons que certaines propriétés des cadres sont également pertinentes pour décrire les réseaux sémantiques qui structurent les paradigmes en dérivation et que nous cherchons à identifier. Nous résumons ici les principaux points en commun.

- les cadres décrivent des réseaux sémantiques qui sont indépendants de leur réalisation syntaxique. De manière similaire, les réseaux sémantiques que nous cherchons à identifier sont également indépendants de leur réalisation dans le lexique (voir Chapitre 1) ;
- les cadres contiennent des rôles interprédictibles. De la même manière, nous cherchons à identifier des réseaux sémantiques contenant des concepts interprédictibles ;
- les rôles sémantiques dans *FrameNet* sont associés à des catégories ontologiques, même si cela n’est pas systématique. Dans notre description des réseaux qui structurent les paradigmes, nous distinguons également l’aspect relationnel (i.e. le rôle) et l’aspect ontologique (restrictions appliquées sur le rôle), comme nous le montrons au Chapitre 4 ;
- les cadres peuvent être reliés par des relations inter-cadre. De manière similaire, nous pouvons établir des relations entre des réseaux qui structurent des paradigmes distincts.

La question qui se pose maintenant est celle de savoir si les cadres tels qu’ils sont présents dans *FrameNet* ou dans sa version française *Asfalda* (Candito *et al.*, 2014) sont directement exploitables pour notre tâche de délimitation des paradigmes. Comme les cadres sont conçus pour décrire l’interface entre sémantique et syntaxe et n’ont pas d’objectifs de description morphologique, nous considérons qu’ils ne sont pas utilisables tels quels pour décrire les relations qui structurent les familles. En revanche, certains réseaux sémantiques que nous allons présenter dans les Chapitres 4 et 5 sont similaires à des cadres que l’on peut trouver dans *FrameNet*.

3.6 Bilan

Nous avons consacré ce chapitre à la notion de rôle sémantique et à la Sémantique des Cadres. Nous avons commencé par montrer des exemples d’usage de rôles sémantiques qui ont été utilisés pour décrire des relations morphosémantiques (Salvadori et Huyghe, 2023). Malgré les nombreux travaux qui ont été consacrés à la description morphosémantique et à la description des propriétés sémantiques des affixes (Booij, 1986; Namer et Villoing, 2008; Rosenberg, 2008; Schnedecker et Aleksandrova, 2016; Huyghe, 2021), à notre connaissance, il n’existe pas un

inventaire de rôles sémantiques qui peut décrire en extension les relations entre membres des familles dérivationnelles.

La Sémantique des Cadres, pour sa part, se base sur des scénarios qui contiennent des rôles sémantiques interprédictibles. Les principes sur lesquels se base cette théorie sont compatibles avec nos besoins de descriptions de réseaux sémantiques qui structurent les familles. Par conséquent, nous considérons que nous pouvons nous inspirer de ce cadre théorique pour réaliser un découpage des familles dérivationnelles en familles paradigmatiques basé sur des scénarios cognitifs similaires aux cadres de *FrameNet* et sur les participants qui interviennent dans ces scénarios cognitifs.

Dans *FrameNet*, les cadres sont présentés au moyen d'histoires abstraites qui mettent en relation en même temps deux ou plusieurs éléments du cadre. Ces histoires révèlent les relations entre les éléments du cadre en le mettant dans un contexte en prose. Comme nous allons voir dans le Chapitre 4, nous prenons le format des descriptions des cadres comme source d'inspiration pour produire des histoires que nous utilisons pour caractériser les relations entre plusieurs membres appartenant à la même famille. Ensuite, nous organisons le contenu régulier de ces histoires dans des objets sémantiques que nous appelons *meaning bundles* : comme les cadres, les *meaning bundles* décrivent des situations prototypiques et leurs participants réguliers. Dans un *meaning bundle*, de la même manière que dans un cadre, les participants sont interprédictibles. Comme nous le montrons dans le Chapitre 4, nous allons utiliser les *meaning bundles* pour délimiter les paradigmes dérivationnels.

Chapitre 4

Une démarche *story-based* pour les paradigmes dérivationnels

Sommaire

4.1	Introduction à notre démarche	90
4.2	Données de départ et structures à produire	92
4.3	Organisation du lexique et structuration des paradigmes	95
4.4	Décrire la sémantique des relations morphologiques	95
4.5	Pertinence des modèles génératifs de langage	98
4.6	Formaliser le contenu des histoires	99
4.7	Variété du contenu sémantique des histoires	104
4.8	Aligner les relations contenues dans les histoires	109
4.9	Subset Meaning bundles	112
4.10	Traitement des autres sous-ensembles de la famille	118
4.11	Meaning bundles au niveau d'une famille	124
4.12	FMBs obtenus à partir d'autres familles et LMBs	127
4.13	Projection des LMBs sur le lexique	137
4.14	Bilan	142

BILAN DU CHAPITRE 1 : La notion de famille dérivationnelle joue un rôle central dans l'organisation du lexique construit. Cependant, la délimitation de cette notion reste un problème

à résoudre dans la littérature (Hathout et Namer, 2019). Des travaux récents considèrent que les familles sont structurées par des réseaux conceptuels (Štekauer, 2014; Roché, 2017b). Au sein du système conceptuel, les concepts sont reliés par des relations de nature variée (taxinomiques, de contiguïté) et s’organisent en réseau. L’interface entre lexique et système conceptuel a été discutée dans plusieurs travaux en sémantique lexicale : les relations lexicales reflètent en partie les réseaux contenus dans le système conceptuel (Cruse *et al.*, 1986; Murphy, 2003).

BILAN DU CHAPITRE 2 : L’organisation paradigmatique des relations dérivationnelles et du lexique construit a été discutée dans plusieurs travaux en morphologie dans les dernières décennies (Van Marle, 1985; Bauer, 1997; Hathout et Namer, 2019). Plusieurs auteurs considèrent que les paradigmes dérivationnels sont structurés par des relations sémantiques récurrentes et se composent de familles paradigmatiques contenant les mêmes relations (Bonami et Strnadová, 2019; Hathout et Namer, 2022). En revanche, la question de la délimitation des paradigmes en dérivation reste irrésolue et est liée à la problématique de la délimitation des familles.

BILAN DU CHAPITRE 3 : La sémantique des cadres (Fillmore *et al.*, 1976; Ruppenhofer *et al.*, 2016) se sert de sortes de scénarios pour l’annotation sémantique de textes et pour la description de l’interface entre rôles sémantiques et leur réalisation syntaxique dans les corpus. Nous proposons d’utiliser des scénarios similaires aux cadres de Fillmore pour délimiter sémantiquement les familles paradigmatiques et les paradigmes en dérivation. Nous établissons un ensemble de correspondances entre les éléments principaux de la sémantique des cadres (cadres, rôles, exemples de corpus, etc.) et réseaux sémantiques que nous cherchons à définir pour délimiter les paradigmes dérivationnels.

4.1 Introduction à notre démarche

Nous proposons dans ce chapitre une démarche théorique et générale pour extraire des familles paradigmatiques sémantiquement délimitées à partir de familles dérivationnelles. Les familles paradigmatiques sont des sous-familles alignables qui peuvent être organisées dans des

paradigmes dérivationnels. Comme nous l’avons indiqué dans le Chapitre 2, nous adoptons la proposition de Antoniova et Štekauer (2016) et Bonami et Strnadová (2019) en considérant les relations sémantiques comme étant premières pour la structuration des paradigmes en dérivation. Notre démarche est guidée par la dimension sémantique des relations morphologiques existantes entre les lexèmes contenus dans les familles.

Nous constatons une première difficulté : nous ne disposons pas de descriptions systématiques des relations existantes entre les lexèmes des familles dérivationnelles contenues dans le lexique construit français. Nous proposons de les acquérir en utilisant ce que nous appelons des “histoires”. Les histoires sont des énoncés construits autour d’un sous-ensemble des lexèmes d’une famille dérivationnelle. Les histoires peuvent être vues comme le résultat d’une tâche d’élicitation dans laquelle on demanderait à des locuteurs de raconter des histoires qui font intervenir les lexèmes appartenant à une famille dérivationnelle.

Nous utilisons les histoires pour identifier les relations sémantiques récurrentes entre les lexèmes des familles et organiser ces relations dans ce que nous appelons des *meaning bundles* (MB). Les MBs que nous proposons sont semblables aux cadres de Fillmore (Fillmore *et al.*, 1976; Fillmore et Baker, 2001) que nous avons introduit au Chapitre 3 : ils décrivent des sortes de situations cognitives dans lesquelles des participants qui sont fortement associés interviennent régulièrement. Le contenu des MBs est une abstraction des relations sémantiques que l’on trouve régulièrement dans les histoires. La relation entre *meaning bundles* et histoires est similaire à celle qui existe entre les cadres de *FrameNet* et les exemples de corpus annotés : les MBs décrivent le contenu régulier des histoires de la même manière que les cadres contiennent des participants et des relations régulièrement présents dans les phrases de corpus annotées (voir Chapitre 3). D’autre part, contrairement aux cadres de Fillmore, nous représentons les MBs sous forme de réseaux sémantiques.

Nous nous servons de trois types distincts de *meaning bundles*, à savoir des *subset meaning bundles* (SMB), de *family meaning bundles* (FMB) et de *lexical meaning bundles* (LMB), que nous allons présenter dans l’illustration de la démarche. Ces trois objets sémantiques se situent sur trois niveaux de généralisation distincts.

La dernière étape de notre démarche consiste à projeter les LMBs sur le lexique construit. Cette étape nous permet d'obtenir des familles paradigmatiques sémantiquement délimitées et qui sont alignables dans des paradigmes.

4.2 Données de départ et structures à produire

Nos données de départ sont les familles dérivationnelles contenues dans le lexique (construit). Comme nous l'avons montré au Chapitre 2, les familles dérivationnelles complètes ne fonctionnent pas en tant que paradigmes. Les paradigmes contiennent des ensembles d'éléments qui sont interprédictibles : en revanche, si on considère une très grande famille comme celle de HERBE que nous avons vue au Chapitre 2, nous pouvons voir qu'elle contient des lexèmes qui ne sont pas sémantiquement connexes et qui ne sont pas prédictibles l'un par rapport à l'autre sur le plan sémantique. Par exemple, la présence de HERBORISTE dans la famille de HERBE ne prédit pas la présence de DÉSHÉRBANT ou de HERBAGÈRE dans la même famille. D'autre part, nous avons vu dans le Chapitre 2 que l'on peut identifier dans la famille de HERBE des morceaux qui contiennent des lexèmes interprédictibles : la présence du verbe DÉSHÉRBER prédit la présence dans la même famille du nom DÉSHÉRBANT, qui dénote un produit qui est utilisé dans le procédé dénoté par le verbe DÉSHÉRBER. Il en va de même pour HERBORISTE, qui est interprédictible sur le plan sémantique par rapport à HERBORISTERIE : HERBORISTE dénote une personne qui vend des herbes médicinales et HERBORISTERIE dénote respectivement l'activité de vente et d'usage d'herbes médicinales et la boutique où l'herboriste travaille.

Cela signifie que **(a)** les familles dérivationnelles complètes ne constituent pas des paradigmes ; **(b)** à l'intérieur des familles dérivationnelles, on a des sous-ensembles qui contiennent des lexèmes interprédictibles qui peuvent constituer des paradigmes. Ces sous-ensembles correspondent à la notion de famille paradigmatique que nous avons introduite au Chapitre 2. En outre, les familles paradigmatiques que nous cherchons contiennent des relations sémantiques récurrentes et peuvent être alignées dans des paradigmes avec d'autres familles paradigmatiques qui contiennent les mêmes relations.

Nous voulons extraire des familles dérivationnelles contenues dans le lexique construit toutes les familles paradigmatiques que nous pouvons identifier. En (29) nous considérons un exemple de famille de départ plus simple que la famille de HERBE pour illustrer cette tâche d'identification. Le lexème POTERIE en (29) dénote trois sens : 'métier du potier', 'objet fabriqué par le potier' et 'atelier du potier'. Pour simplifier l'illustration, dans la famille en (29) nous considérons uniquement le sens 'métier du potier'. En outre, dans la famille en (29) nous ne considérons pas REMPOTAGE, car il dénote le même procès que rempoter. La relation entre verbe et nom d'action dénotant le même procédé est abordée dans le Chapitre 6.

$$(29) \quad F1 = \{\text{POT.N, POTIER.N, POTERIE.N, REMPOTER.V, REMPOTEUR.N}\}$$

Dans la famille (29), POT, POTIER et POTERIE sont sémantiquement connexes : un potier fait des pots et la poterie est l'activité de fabrication des pots et le métier du potier. Nous pouvons considérer que POT, POTIER et POTERIE sont sémantiquement interprédictibles et peuvent faire partie de la même famille paradigmatique. Inversement, un potier n'intervient pas à priori dans l'action de rempoter : il s'occupe de fabriquer les pots, mais pas de les vider ni de les remplir pour y mettre des plantes. La même chose vaut pour POTERIE : l'activité de faire des pots n'a rien à voir avec l'action de rempoter. Cela signifie que ni POTIER ni POTERIE sont dans la même famille paradigmatique de REMPOTER. À l'inverse, REMPOTER, POT et REMPOTEUR sont sémantiquement connexes, car l'action de rempoter implique l'utilisation d'un pot pour transférer une plante et cette action est généralement réalisée par un agent humain. On peut considérer donc que POT, REMPOTER et REMPOTEUR sont sémantiquement interprédictibles et qu'ils font partie de la même famille paradigmatique. La famille $F1$ peut donc être divisée en deux familles paradigmatiques f_1 et f_2 , qui sont présentées en (30).

$$(30) \quad \begin{array}{l} \text{a. } f_1 = \{\text{POT.N, POTIER.N, POTERIE.N}\} \\ \text{b. } f_2 = \{\text{POT.N, REMPOTER.V, REMPOTEUR.N}\} \end{array}$$

Les familles paradigmatiques que nous voulons produire sont alignées dans des paradigmes dérivationnels. Par conséquent, on peut trouver d'autres familles paradigmatiques qui contiennent

les mêmes relations sémantiques que celles en (30). Pour réaliser l’alignement, nous nous basons sur des critères ontologiques et relationnels. En (31) nous fournissons deux exemples de familles paradigmatiques alignables avec celles données en (30).

- (31) a. $f_3 = \{\text{BOTTE.N, BOTTIER.N, BOTTERIE.N}\}$
 b. $f_4 = \{\text{BOUTEILLE.N, EMBOUTEILLER.V, EMBOUTEILLEUR.N}\}$

La famille paradigmatique en (31a) contient un nom d’artefact (BOTTE), le nom dénotant le fabricant de bottes (BOTTIER) et l’activité de fabrication des bottes (BOTTERIE). BOTTERIE est considérable comme un nom de domaine d’activité, car il est compatible avec la tournure “ N_{humain} a reçu une formation dans le domaine de la botterie” (Haas et al., 2023, p.62). Nous précisons que BOTTERIE dénote également le lieu où les bottes sont fabriquées, mais à ce stade de l’illustration nous nous concentrons sur son sens de ‘domaine d’activité’.

Nous retrouvons dans (31a) les mêmes relations qui existent entre les lexèmes en (30a) : le bottier fabrique des bottes, la botterie et l’activité de fabrication des bottes et la botterie est l’activité principale du bottier. Sur cette base, nous considérons que f_3 est dans le même paradigme qui contient f_1 .

Pour sa part, la famille paradigmatique en (31b), comme celle en (30b), contient un nom dénotant un récipient (BOUTEILLE), un verbe qui réfère à l’action de déplacer une entité dans ce récipient (EMBOUTEILLER) et un nom d’une personne qui réalise de manière intentionnelle ce processus (EMBOUTEILLEUR). Sur cette base, nous considérons que f_2 est dans un paradigme dérivationnel qui contient également f_4 .

Nous définissons notre tâche comme suit (32).

- (32) **Définition de la tâche** : Soit F une famille dérivationnelle, trouver tous les paradigmes qui contiennent des sous-ensembles de lexèmes de F .

Notre objectif est donc d’identifier les familles paradigmatiques contenues dans F et les paradigmes dérivationnels qui contiennent ces familles paradigmatiques.

4.3 Organisation du lexique et structuration des paradigmes

Nous nous intéressons à la structuration du lexique par les relations morphologiques, à savoir au lexique construit. Le lexique construit contient des familles dérivationnelles et chaque famille dérivationnelle F est représentable comme un graphe, comme nous l'avons vu dans le Chapitre 1. Les sommets de ce graphe représentent les lexèmes contenus dans la famille et les arcs représentent les relations morphologiques qui existent entre ces lexèmes. Les relations morphologiques qui relient les lexèmes dans une famille dérivationnelle sont des relations de partage de forme et de sens (Haspelmath et Sims, 2013, p.2). Comme nous l'avons indiqué dans les chapitres précédents, nous faisons l'hypothèse que les relations morphologiques contenues dans les familles dérivationnelles du lexique sont paradigmatiques, dans le sens où elles peuvent être alignées dans des paradigmes (33).

(33) **Hypothèse₁** : le lexique construit est structuré par des paradigmes.

Nous considérons que la dimension sémantique des relations morphologiques détermine la structure des paradigmes dérivationnels (Antoniova et Štekauer, 2016; Bonami et Strnadová, 2019). Deux relations entre les membres de deux couples de lexèmes de deux familles paradigmatiques différentes sont alignables si la relation sémantique qui existe à l'intérieur du premier couple de lexèmes est la même que celle qui existe dans le deuxième couple de lexèmes (Bonami et Strnadová, 2019, p.4). Cette notion d'alignement est utilisée pour construire le paradigme en Figure 4.1 que nous avons introduit dans le Chapitre 2 : comme on peut le voir, les relations qui structurent les couples {LAVÉ.V, LAVAGE.N}, {FORMER.V, FORMATION.N} et {GONFLER.V, GONFLEMENT.N} sont alignées sur la base de leur contenu sémantique. Les différences au niveau formel sont secondaires pour la structuration des paradigmes que nous cherchons à produire.

4.4 Décrire la sémantique des relations morphologiques

Notre objectif dans ce chapitre est de mettre au point une méthode permettant d'aligner dans des paradigmes les familles paradigmatiques qui contiennent les mêmes relations sémantiques.

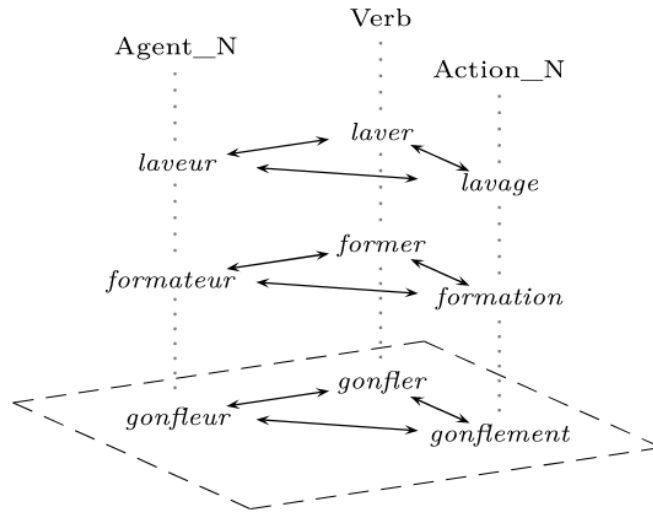


FIGURE 4.1 – Paradigme dérivationnel structuré par des relations sémantiques (Bonami et Strnadová, 2019, p.172).

Il existe plusieurs ressources qui contiennent des descriptions de relations lexicales, comme par exemple *Wordnet* (Miller, 1998; Fellbaum, 2010) : en revanche cette ressource présente un inventaire trop limité pour être utilisable dans notre tâche (voir Chapitre 1). Sa version française, le *WOLF* (Sagot et Fišer, 2008) ou encore le réseau lexical *JeuxDeMots* (Lafourcade et Joubert, 2008) ne sont pas utilisables non plus, car leurs inventaires ne répondent pas à notre besoin de décrire de manière extensive les relations dérivationnelles. Les ressources basées sur la sémantique des cadres comme *FrameNet* ou *Asfalda* (Ruppenhofer *et al.*, 2016; Candito *et al.*, 2014), que nous avons présentées en Chapitre 3, ne sont directement exploitables non plus, car leur description est centrée sur l’interface entre sémantique et syntaxe. Elles ne couvrent pas suffisamment les relations sémantiques pertinentes pour la morphologie dérivationnelle et la structuration du lexique construit.

Sur cette base, nous avons deux problèmes : **(a)** nous ne connaissons pas la totalité des relations sémantiques qui existent dans le lexique construit ; **(b)** nous ne savons pas déterminer si deux relations dérivationnelles peuvent être alignées sur le plan sémantique. Au vu du manque de ressources directement exploitables pour résoudre ce double problème, nous avons

choisi d’adopter une approche qui s’inspire à la fois d’une approche distributionnelle de mise en contexte des lexèmes (Harris, 1954) et des situations prototypiques des cadres de Fillmore (Fillmore *et al.*, 1976). Nous proposons de caractériser les relations sémantiques dans un sous-ensemble de lexèmes au moyen d’histoires qui placent ces lexèmes dans un contexte naturel. Un des avantages des histoires est qu’elles permettent de décrire plusieurs relations entre lexèmes en même temps. Un exemple est fourni en (34). Ce qui nous intéresse des histoires est leur contenu régulier en termes de participants et de relations. Pour chaque histoire, nous identifions les participants en soulignant les parties de l’histoire qui y font référence. Parmi les participants, nous incluons également des expressions temporelles, de manière similaire à ce qui est fait dans *FrameNet* (Ruppenhofer *et al.*, 2016), où l’ancrage temporel est considéré comme étant un participant non-noyau des cadres (voir Chapitre 3).

Par exemple, nous considérons que l’histoire en (34) implique les participants suivants : le potier, l’art de la poterie, son fils, les pots, l’argile, l’atelier et l’expression d’ancrage temporel *dans les dernières semaines*.

- (34) *Dans les dernières semaines, le potier a enseigné l’art de la poterie à son fils. Ils ont créé de superbes pots en argile dans son atelier.*

Nous aurions idéalement souhaité extraire d’un corpus des contextes similaires à l’histoire en (34) dans lesquels les lexèmes contenus dans les familles apparaissent. Toutefois, les corpus français de grande taille extraits du web comme *FrWac* (Baroni *et al.*, 2009) ou *FrCow* (Schäfer et Bildhauer, 2012; Schäfer, 2015) ne contiennent pas de manière systématique des histoires comme celle en (34). Une solution de rechange serait de mettre en place une campagne de recueil d’histoires auprès de locuteurs du français pour éliciter ces descriptions des relations sémantiques. Malheureusement, la quantité de relations à décrire, de descriptions à produire et le nombre de locuteurs à recruter rendent impossible une collecte à grande échelle d’histoires produites par des locuteurs. Sur ce fait, nous allons illustrer la démarche que nous proposons en utilisant des histoires que nous avons créées nous-mêmes. Une grande partie des histoires

présentées dans cette thèse a été produite en utilisant *ChatGPT*, l’agent conversationnel de *OpenAI*. Nous motivons ce choix en Section 4.5.

Ces histoires sont un outil heuristique que nous utilisons pour expliciter les relations qui existent entre les lexèmes des familles et pour mettre en œuvre la démarche que nous proposons. Elles représentent une double approximation : **(a)** nous considérons que les histoires explicitent et motivent les relations sémantiques entre les lexèmes ; **(b)** nous considérons que les histoires que nous produisons contiennent des descriptions similaires à celles qui sont produites par des locuteurs.

4.5 Pertinence des modèles génératifs de langage

Il existe désormais des modèles d’IA génératifs qui sont capables de générer des histoires très proches de celles que nous visons¹². Une partie des histoires que nous présentons dans cette thèse ont été produites en utilisant *chatGPT*, l’agent conversationnel de OpenAI¹³. La version gratuite de cet agent conversationnel se sert du modèle *GPT3.5* (Brown *et al.*, 2020). Il s’agit d’un modèle basé sur des réseaux de neurones, qui a été préentraîné sur une grande quantité de données multilingue provenant de sources variées (*CommonCrawl*, *Wikipedia* et plein d’autres). Comme d’autres modèles de langage basés sur des réseaux de neurones, *GPT* a la capacité de reconnaître la proximité sémantique entre les mots. Il produit des conversations, des réponses et des histoires sémantiquement cohérentes et très proches de celles qu’un locuteur humain pourrait produire. Pour donner un exemple, l’histoire en (35) a été générée par *ChatGPT* à partir du prompt suivant : « *Raconte une histoire de 100 mots au maximum qui contient les mots “pot” et “potier”.* ».

(35) *Il était une fois un potier nommé Marc, qui aimait créer des pots en argile dans son atelier. Un jour, il créa un magnifique pot, parfaitement équilibré et décoré avec des*

12. En revanche, quand nous avons commencé à travailler sur cette thèse en octobre 2018, la disponibilité et performance des modèles génératifs de langage était beaucoup plus réduite.

13. <https://chat.openai.com/>

*motifs complexes. Ce pot devint très populaire parmi les clients de Marc, qui ne taris-
saient pas d'éloges sur son travail. Et c'est ainsi que Marc devint célèbre pour la qualité
de son travail, en particulier pour ses pots exquis.*

Tout en étant conscients des limites et des biais qui caractérisent les modèles langagiers de grande taille comme *GPT* (Lucy et Bamman, 2021; Bender *et al.*, 2021) et de l'importance du prompt donné comme input au modèle, nous considérons que les histoires produites par *ChatGPT* sont pertinentes pour être utilisées dans notre démarche. Leur intérêt principal est le fait d'être des histoires "prototypiques" qui font intervenir des participants qui apparaissent souvent dans le même contexte linguistique. En outre, nous considérons que ce type de modèle pourrait être utilisé pour un futur recueil de données à grande échelle.

4.6 Formaliser le contenu des histoires

Au début de notre démarche, nous ne connaissons pas les sous-ensembles des lexèmes d'une famille dérivationnelle qui correspondent à des familles paradigmatiques. Par conséquent, nous avons besoin de tester tous les sous-ensembles de la famille de départ pour établir lesquels correspondent à des familles paradigmatiques. Nous choisissons d'utiliser une méthode de "force brute" en partant d'une famille comme celle de POT en (29) et en considérant tous ses sous-ensembles de cardinalité ≥ 2 . Ces sous-ensembles sont contenus dans un ensemble que nous appelons $cov(F1)$ et qui a les propriétés d'un recouvrement (36), c'est-à-dire que les sous-ensembles d'une famille dérivationnelle peuvent partager des lexèmes et se chevaucher : POT, par exemple, apparaît dans plusieurs sous-ensembles en (36).

$$(36) \quad cov(F1) = \{ \{ POT.N, POTERIE.N \}, \{ POT.N, POTIER.N \}, \dots, \{ REMPOTER.V, REMPOTEUR.N \}, \dots, \{ REMPOTER.V, POTERIE.N \}, \{ POT.N, POTIER.N, POTERIE.N \}, \dots, \{ POT.N, POTERIE.N, POTIER.N, REMPOTER.V, REMPOTEUR.N \} \}$$

Pour chaque sous-ensemble contenu en (36), nous produisons un ensemble d'histoires qui contiennent des instances de ses lexèmes. Une histoire peut être formée par une seule phrase,

comme celle en (37) ou par plusieurs phrases, comme dans l'histoire présentée précédemment en (35).

(37) *Un jour, un potier nommé Marc créa un magnifique pot dans son atelier.*

Nous faisons l'hypothèse que les concepts réalisés par les lexèmes qui sont fortement associés dans les familles dérivationnelles vont être instanciés ensemble dans la plupart des histoires que l'on peut produire. Par exemple, nous nous attendons à ce que dans une histoire construite à partir du sous-ensemble {POT.N, POTIER.N} on parle également d'une activité ou d'une action de fabrication d'un pot, de matériaux utilisés et d'un lieu où le pot est fabriqué, comme dans l'histoire en (35). De manière similaire, si on construit une histoire à partir de {POT.N, POTERIE.N} on s'attend à trouver dans cette même histoire une personne qui fabrique le pot, comme dans l'histoire en (38) construite sur ce couple {POT.N, POTERIE.N}.

(38) *Claire, animée par sa passion pour la poterie, passait des heures à façonner l'argile pour créer de merveilleux pots. Son atelier était rempli de pots de toutes formes et tailles, des vases élégants aux bols pratiques. La poterie était pour lui un moyen d'exprimer sa créativité et de capturer la beauté de la nature dans un simple morceau d'argile.*

D'autre part, nous nous attendons à ce que les concepts lexicalisés par les lexèmes qui appartiennent à la même famille mais qui ne sont pas sémantiquement connexes apparaissent ensemble dans peu d'histoires et que leur cooccurrence dans ces histoires soit accessoire. Par exemple, l'histoire en (39), construite autour du sous-ensemble {POT.N, REMPOTER.V} pourrait difficilement impliquer une activité de fabrication des pots ou des personnes qui créent des pots. Cette histoire décrit le déplacement d'une plante dans un pot et n'a rien à voir avec la fabrication des pots : elle implique des participants qui sont sémantiquement cohérents avec une action de rempotage, comme des plantes, des nutriments pour les nourrir, une personne qui fait le rempotage, etc.

(39) *Sophie aimait collectionner des plantes et avait une grande variété de pots. Un jour, elle décida de rempoter toutes ses plantes dans de nouveaux pots plus grands pour leur*

donner plus d'espace pour grandir. Elle passa toute la journée à rempoter ses plantes et à les nourrir avec des nutriments pour favoriser leur croissance. Le lendemain matin, en ouvrant les stores de son balcon, elle remarqua que toutes ses plantes semblaient plus heureuses dans leurs nouveaux pots. Elle se dit que le rempotage avait été une bonne décision et qu'elle devrait le faire plus souvent pour prendre soin de ses plantes.

- (40) **hypothèse₂** : Les concepts lexicalisés par des lexèmes sémantiquement connexes dans une famille dérivationnelle vont être régulièrement instanciés dans les mêmes histoires. Au contraire, les concepts lexicalisés par les lexèmes d'une famille qui ne sont pas connexes sur le plan sémantique, vont apparaître ensemble plus rarement dans les histoires et leur cooccurrence sera accessoire.

Nous utilisons les histoires pour identifier les relations sémantiques récurrentes entre les lexèmes des familles dérivationnelles. Cela se base sur l'hypothèse selon laquelle nous sommes en mesure d'identifier des relations sémantiques paradigmatiques dans le lexique sur la base des relations distributionnelles régulières contenues dans les histoires. Autrement dit, nous considérons que les propriétés sémantiques d'une unité lexicale sont étroitement reliées à son environnement linguistique (Harris, 1954), qui dans notre cas est le contenu des histoires.

- (41) **hypothèse₃** : Nous pouvons identifier les relations sémantiques paradigmatiques dans le lexique sur la base de la distribution des relations contenues dans les histoires.

Pour identifier les relations sémantiques régulièrement présentes dans les histoires, nous proposons de les formaliser au moyen d'objets plus faciles à manipuler que du texte. Nous utilisons des graphes sémantiques similaires aux réseaux sémantiques que nous avons introduits au Chapitre 1 (Quillan, 1966; Sowa, 2014). Nous choisissons d'utiliser des graphes, à savoir des structures en réseau, car les familles dérivationnelles sont également représentées au moyen de réseaux et ce choix nous permet d'utiliser des réseaux pour représenter à la fois les familles dérivationnelles et les histoires. Les histoires que nous représentons avec nos graphes sémantiques décrivent des référents tels que des entités, des actions et des propriétés, à savoir les trois catégories sémantiques principales identifiées par Croft (1991). La représentation sous forme

de graphe de l’histoire est le résultat d’une analyse de l’histoire elle-même. Cette analyse est effectuée de manière séparée pour toutes les histoires produites à partir de chaque sous-ensemble de lexèmes qui est contenu dans une famille dérivationnelle.

Pour illustrer la formalisation du contenu des histoires, nous considérons un premier sous-ensemble {POT.N, POTIER.N} et reprenons l’histoire en (42) que nous avons présentée précédemment.

(42) Un jour, un potier nommé Marc créa un magnifique pot dans son atelier.

Pour simplifier la représentation du contenu des histoires comme celle en (42), dans nos graphes sémantiques nous nous inspirons de la proposition de Davidson (1967, 1969) en reliant les participants de l’histoire au moyen de l’événement dans lequel ils sont impliqués. Selon Davidson (1967), les événements sont des entités spatio-temporelles introduites par des prédicats qui font référence à des actions délimitées dans le temps et dans l’espace (Maienborn, 2011, p.1). Dans nos graphes, nous représentons les actions et les événements que les histoires introduisent dans le même sommet (i.e. un double cercle), comme nous le faisons dans la Figure 4.2. Nous représentons les relations entre l’événement et ses participants au moyen de rôles sémantiques, comme cela est fait dans plusieurs travaux néo-davidsoniens (Parsons, 1995; Higginbotham, 2000, entre autres). Cependant, nous nous situons à un niveau de généralisation inférieur par rapport aux rôles sémantiques que nous avons présentés dans le Chapitre 3. Nous faisons ce choix, car des rôles sémantiques tels qu’AGENT, PATIENT ou INSTRUMENT¹⁴ sont trop généraux pour identifier des relations sémantiques et des paradigmes dérivationnels stables (voir Chapitre 2). Pour cette raison, nous utilisons des rôles sémantiques plus spécifiques tels que FABRICANT, OBJET FABRIQUÉ et LIEU DE FABRICATION. Le niveau de généralisation de ces rôles est similaire à celui utilisé par les cadres de Fillmore que nous avons traités au Chapitre 3.

Contrairement aux catégories ontologiques, qui sont généralement attribuées au moyen de tests, les rôles sémantiques sont attribués sur la base de leur conformité à une définition (Salvadori et Huyghe, 2023, p.11). Cela est fait aussi dans *FrameNet* où chaque participant d’un cadre

14. Nous utilisons le majuscule pour noter les rôles sémantiques, comme nous l’avons déjà fait au Chapitre 3.

est accompagné par une définition (voir Chapitre 3). Sur cette base, nous considérons que :

- LIEU DE FABRICATION correspond à tout lieu physique dans lequel une activité ou une action de fabrication se déroule ;
- OBJET FABRIQUÉ correspond à tout artefact qui est le résultat d'un procédé de fabrication ;
- FABRICANT correspond à un agent humain qui volontairement fabrique un artefact.

Nous attribuons ces rôles aux participants de l'histoire (42) : le potier joue le rôle du FABRICANT, le pot joue le rôle d'OBJET FABRIQUÉ et l'atelier est le LIEU DE FABRICATION. Le graphe sémantique que nous proposons pour l'histoire (42) se structure en étoile autour de l'événement décrit par *créer*, comme on peut le voir en Figure 4.2. L'action et l'événement qu'elle introduit fonctionnent comme une sorte de pivot sémantique de l'histoire représentée en Figure 4.2 et les participants de l'histoire sont reliés à l'événement par des relations qui représentent leur rôle sémantique. Nous typons la relation temporelle entre l'événement et son ancrage temporel en utilisant le rôle DATE, qui dénote le jour où une action donnée se déroule. En outre, nous considérons que le potier est relié par une relation d'identité à *Marc*, car il s'agit de la même personne.

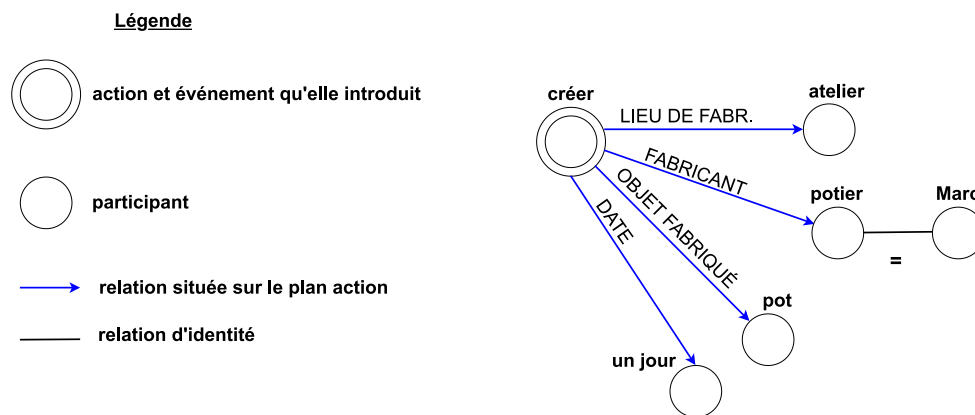


FIGURE 4.2 – Le contenu sémantique de l'histoire (42) est représenté au moyen d'un graphe. Nous représentons les actions ancrées dans le temps et dans l'espace et les événements qu'elles introduisent au moyen d'un double cercle et nous relierons l'événement à ses participants au moyen d'arcs orientés en bleu, qui correspondent à des relations thématiques.

Nous représentons les relations entre l'événement et ses participants au moyen d'arcs orientés en bleu et nous considérons que ces relations se situent sur le même plan que le *réseau action* proposé par Roché (2017b). Dans les graphes, les relations thématiques sont orientées de l'événement vers ses participants car nous considérons que c'est l'événement qui attribut aux participants leur rôle dans ce dernier.

4.7 Variété du contenu sémantique des histoires

Dans la section précédente nous avons illustré la formalisation du contenu d'une histoire simplifiée sous forme de graphe sémantique. Cependant, des véritables histoires produites à partir de sous-ensembles de lexèmes contiennent à priori beaucoup plus d'éléments au niveau sémantique que l'histoire en (42). Ces éléments sont de nature sémantique variée et ne se limitent pas à des actions et à des participants. En (43) nous fournissons un exemple d'histoire plus complexe construite sur le même sous-ensemble utilisé précédemment {POT.N, POTIER.N}.

- (43) *Dans un petit village, un potier nommé Pierre était connu pour ses créations uniques. Chaque jour, il se rendait à son atelier et façonnait des pots magnifiques. Les habitants venaient de loin pour acheter ses œuvres d'art en argile, et Pierre était fier de savoir que ses pots apportaient de la joie aux gens du village.*

L'histoire en (43) diffère de l'histoire que nous avons représentée en Section 4.6 car elle ne se structure pas autour d'une action, mais autour de plusieurs ensembles organisés d'actions, à savoir des activités (dans le sens du terme utilisé par Roché (2017b)). Par exemple, elle décrit l'activité de création de pots du potier, l'activité de vente de ces pots et le déplacement régulier des habitants du village pour acheter ces pots. Ces activités décrivent des procès réguliers, comme l'indique le marqueur d'habitualité *chaque jour*. La distinction entre action et activité proposée par Roché (2017b) est similaire à la distinction entre prédicats *stage level* et *individual level* proposés par Carlson (1977) : les actions introduisent des événements ponctuels qui sont délimités dans le temps et dans l'espace et que l'on peut considérer comme transitoires, alors que

les activités décrivent des ensembles organisés de plusieurs procès ou une répétition du même procès de manière régulière. Les activités, contrairement aux actions, ne sont pas délimitées temporellement.

Une représentation simplifiée du contenu sémantique de l'histoire (43) est proposée en Figure 4.3. En premier lieu, nous représentons l'activité de création de pots que le potier réalise. Les activités impliquent des participants qui, contrairement aux participants d'une action, sont des participants habituels : le potier (i.e. un FABRICANT habituel) crée des pots magnifiques (i.e. un OBJET FABRIQUÉ habituel) dans son atelier (un LIEU DE FABRICATION habituel). Cette histoire implique également un MATÉRIAU utilisé de manière habituelle dans la fabrication des pots (i.e. l'argile).

En outre, en Figure 4.3 nous représentons l'activité de vente des pots, dans laquelle les habitants sont impliqués en tant qu'ACHETEUR, le potier en tant que VENDEUR et les pots en tant que BIEN COMMERCIALISÉ. Les rôles que nous associons à l'activité de vente sont également présents dans le scénario de transaction commerciale évoqué par (Petrucci, 1996, p.1) et décrit dans *FrameNet*.

En dernier lieu, nous représentons l'activité régulière de déplacement de la part des habitants vers l'atelier pour aller acheter les pots : nous attribuons aux habitants le rôle du VOYAGEUR. Cette étiquette est attribuée à tout participant humain qui réalise de manière intentionnelle un déplacement de moyenne ou longue distance. En ce qui concerne l'activité de déplacement des habitants, nous considérons que l'atelier joue le rôle de la DESTINATION, qui fait partie des rôles sémantiques que nous avons présentés au Chapitre 3.

Nous précisons que nous avons omis en Figure 4.3 la représentation d'autres éléments évoqués par l'histoire que nous considérons comme étant accessoires (ex. la joie apportée aux clients, le fait que les habitants viennent de loin, le fait que le potier est relié par une relation d'identité avec *Pierre*) pour simplifier la représentation. La Figure 4.3 montre que diverses activités peuvent partager des participants : par exemple, le potier est le FABRICANT par rapport à son activité de fabrication et le VENDEUR par rapport à son activité de vente. Nous considérons que les histoires que nous utilisons, tout en décrivant des activités ou des actions diverses, sont toujours

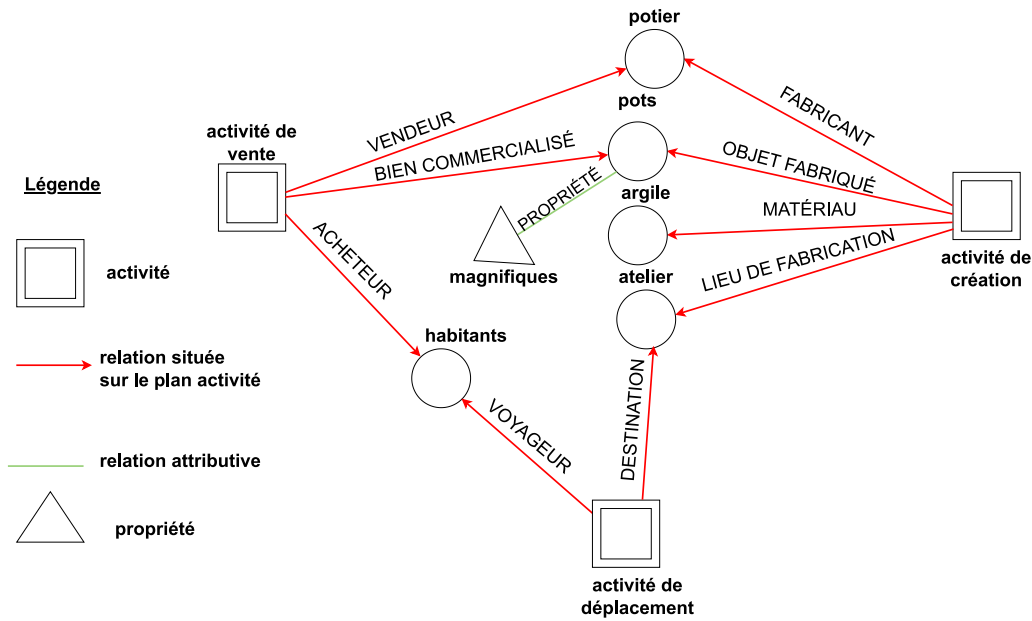


FIGURE 4.3 – Représentation du contenu sémantique de l’histoire en (43). Nous considérons que cette histoire décrit un ensemble d’activités qui coexistent au sein de la même histoire. Nous représentons les activités au moyen de doubles carrés. Nous orientons les arcs de l’activité vers ses participants habituels car nous considérons que c’est l’activité qui attribue aux participants leur rôle dans celle-ci.

représentables au moyen d’un graphe connexe. Dans l’histoire représentée en Figure 4.3, la connexion entre les différentes activités se fait au niveau des participants qui sont impliqués dans plusieurs activités et permettent à l’histoire d’être cohérente sur le plan narratif.

Il existe également des histoires dans lesquelles coexistent des activités et des actions délimitées d’un point de vue temporel, comme l’histoire en (44) que nous avons évoquée précédemment en Section 4.6.

- (44) *Il était une fois un potier nommé Marc, qui aimait créer des pots en argile dans son atelier. Un jour, il créa un magnifique pot, parfaitement équilibré et décoré avec des motifs complexes. Ce pot devint très populaire parmi les clients de Marc, qui ne tarissaient pas d’éloges sur son travail. Et c’est ainsi que Marc devint célèbre pour la qualité de son travail, en particulier pour ses pots exquis.*

Nous proposons une représentation simplifiée du contenu sémantique de l’histoire en (44)

dans la Figure 4.4. Dans cette figure, nous traitons de manière séparée l’activité de création des pots du potier et l’action ponctuelle de création d’un pot parfaitement équilibré et décoré en adoptant la distinction action vs activité proposée par Roché (2017b).

En premier lieu, en Figure 4.4 nous représentons l’activité de fabrication de pots. Par rapport à cette activité, le potier est le FABRICANT habituel, les pots sont l’OBJET FABRIQUÉ habituel et l’argile est le MATÉRIAU utilisé habituellement. En deuxième lieu, nous représentons l’activité de vente de ces pots, car l’histoire parle de clients. Par rapport à cette activité, le potier est le VENDEUR, les pots jouent le rôle du BIEN et les clients jouent le rôle d’ACHETEUR. En outre, nous représentons l’action ponctuelle de fabrication du pot parfaitement équilibré. Par rapport à cette action ponctuelle, le potier est le FABRICANT et le pot est l’OBJET FABRIQUÉ. Nous ajoutons également les motifs complexes dont l’histoire (44) parle, auxquels nous attribuons le rôle de DÉCORATION, qui réfère à toute décoration réalisée sur un produit. Enfin, nous représentons les propriétés associées aux participants dont l’histoire parle, qui sont symbolisées par des triangles.

En Figure 4.4 nous avons choisi de représenter les relations situées sur le plan action et sur le plan activité de manière séparée. Nous avons vu dans le Chapitre 2 que les réseaux action et activité proposés par Roché se situent à la croisée entre des réseaux onomasiologiques (i.e. conceptuels) et des réseaux purement morphologiques. Cependant, lorsque nous traitons des histoires, nous nous plaçons à un niveau purement sémantique et sur le plan sémantique Roché définit une activité comme étant une organisation de plusieurs procès de nature différente ou au moins la répétition systématique du même procès (Roché, 2017b, p.9). Par exemple, les noms dénotant des domaines d’activité comme POTERIE ou BOTTERIE et qui sont compatibles avec la tournure “*faire de la X*” ou “*N_{humain} a reçu une formation dans le domaine de la X*” dénotent des “réitérations” régulières du même procédé (i.e. fabriquer des pots, faire des bottes).

Si nous revenons à l’histoire en (44), nous voyons qu’elle décrit une répétition systématique de la part du potier Marc d’un procès de création de pots, à savoir une activité de fabrication de pots à laquelle il se consacre tous les jours. Cette même histoire ensuite décrit une action ponctuelle dans laquelle le potier Marc est impliqué dans un procès de la même nature, car il

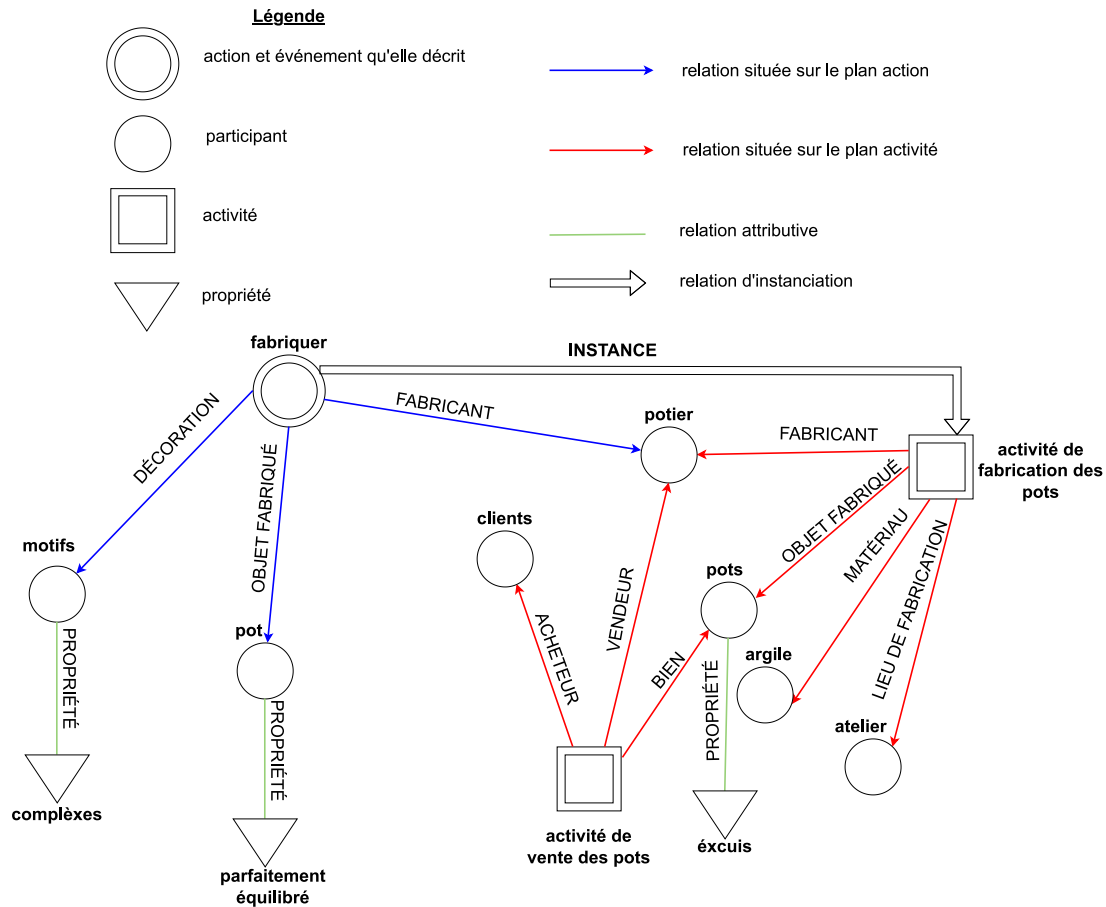


FIGURE 4.4 – Représentation sous forme de graphe sémantique de l’histoire en (44). Nous considérons que dans cette histoire coexistent des relations qui se situent sur le plan action et des relations qui se situent sur le plan activité. Nous considérons que l’action ponctuelle de fabrication de pot est une instance de l’activité de fabrication d’un pot exercée par Marc (nous détaillons ce choix dans la suite de cette section). Cette relation est représentée par une double flèche qui est orientée vers l’activité pour signaler que c’est l’action qui instancie l’activité.

fabrique un pot. Sur cette base, nous considérons que ce que Roché (2017a) appelle une “activité” peut être vu comme une répétition systématique d’un événement décrit par un réseau action et que les deux réseaux ont la même structure au niveau sémantique en termes de participants. Par conséquent, le réseau action peut être vu d’un point de vue sémantique comme une instance d’un réseau activité de même nature. C’est pour cette raison qu’en Figure 4.4 nous relierons le sommet correspondant à l’action de fabrication à l’activité au moyen d’une double flèche qui

représente un lien d’instanciation.

Pour récapituler, les histoires que nous créons à partir d’un sous-ensemble comme {POT.N, POTIER.N} peuvent introduire des actions ponctuelles, des activités ou les deux. Lorsqu’une action et une activité de même nature sont présentes dans la même histoire, elles peuvent être reliées au moyen d’une relation d’instanciation que nous représentons dans nos graphes avec une double flèche orientée de l’action vers l’activité à laquelle nous attribuons l’étiquette INSTANCE.

4.8 Aligner les relations contenues dans les histoires

Dans notre démarche, nous avons besoin d’aligner des relations sémantiques contenues dans des différentes histoires (donc dans des différents graphes) de manière à identifier celles qui sont récurrentes. Les relations récurrentes sont contenues dans des objets que nous appelons des *meaning bundles* (MB). Les MBs sont des objets purement sémantiques qui contiennent une abstraction des concepts et des relations qui coexistent régulièrement dans les histoires. La relation entre *meaning bundles* et histoires est similaire à celle qui existe entre les cadres de *FrameNet* et les phrases de corpus annotées : les MBs ne contiennent que les concepts et relations réguliers dans les histoires et filtrent tout ce qui est épisodique. De manière similaire, les cadres de *FrameNet* que nous avons introduits en Chapitre 3 sont des scénarios abstraits qui ne contiennent que les participants régulièrement instanciés ensemble dans les exemples de corpus annotés (Ruppenhofer *et al.*, 2016).

Définition 8

Meaning bundle

Un *meaning bundle* (MB) est un objet sémantique contenant une abstraction des relations sémantiques et des concepts récurrents dans les histoires créées à partir de collections de lexèmes. Il est semblable à la notion de cadre utilisée par Fillmore (Fillmore *et al.*, 1976; Fillmore et Baker, 2001; Ruppenhofer *et al.*, 2016).

Cependant, à la différence de ce qui est fait pour les cadres de Fillmore, nous représentons les *meaning bundles* sous forme de réseaux¹⁵. Cela nous permet de formaliser les trois objets principaux de notre démarche (i.e. familles dérivationnelles, histoires, MBs) en utilisant des graphes.

Nous précisons que dans notre démarche nous nous servons de trois types distincts de *meaning bundles* :

- *subset meaning bundles* (SMB) : ils réalisent une abstraction au niveau d’un sous-ensemble de lexèmes d’une famille dérivationnelle ;
- *family meaning bundles* (FMB) : ils réalisent une abstraction au niveau de tous les sous-ensembles inscrits dans une même famille dérivationnelle ;
- *lexical meaning bundles* (LMB) : ils réalisent une abstraction au niveau de toutes les familles contenues dans le lexique construit.

Les *meaning bundles* sont constitués à partir d’un grand nombre d’histoires, car nous avons besoin d’identifier les relations récurrentes qui reviennent systématiquement. Sur cette base, nous devons identifier des critères qui permettent d’aligner des relations contenues dans deux ou plusieurs graphes représentant des histoires différentes. Pour illustrer cette étape, nous re-proposons la première histoire simplifiée que nous avons considérée pour le couple {POT.N, POTIER} en Section 4.6 (45a) et nous la comparons avec une deuxième histoire en (45b) produite à partir du même sous-ensemble. Nous représentons l’histoire (45b) avec le graphe proposé en Figure 4.5.

- (45) a. Un jour, un potier nommé Marc créa un magnifique pot dans son atelier.
- b. Un jour, le potier Lucas fabriqua un pot en céramique particulièrement élégant.

Pour aligner deux relations appartenant à deux graphes représentant deux histoires différentes, deux dimensions sémantiques nous semblent être pertinentes : une dimension ontologique et une dimension relationnelle. Pour illustrer ces deux dimensions, en Figure 4.6 nous considérons deux relations qui sont inscrites dans les deux histoires considérées en (45) : la relation

¹⁵. Au contraire, nous avons vu dans le Chapitre 3 que les cadres de *FrameNet* sont décrits par des scénarios abstraits en prose.

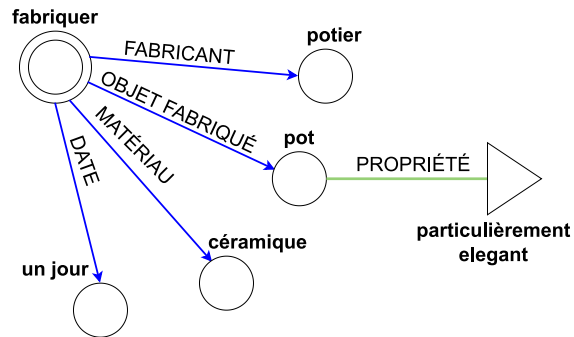


FIGURE 4.5 – Représentation du contenu sémantique de l’histoire en (45b). Dans cette histoire, le potier joue le rôle du FABRICANT, le pot de l’OBJET FABRIQUÉ et la céramique est du MATÉRIAU utilisé.

qui relie l’action de fabrication de l’histoire (45a) au potier Marc et la relation qui relie l’action de fabrication exprimée par l’histoire (45b) au potier Lucas. Nous considérons que ces deux relations peuvent être alignées car (a) sur le plan relationnel, il s’agit de deux relations de type FABRICANT (b) sur le plan ontologique, les actions réalisées sont deux actions de fabrication (c) toujours sur le plan ontologique, les fabricants de ces deux actions sont des humains. Les deux relations concernées sont présentées en Figure 4.6. Nous précisons à nouveau que dans le deux cas le potier est relié par une relation d’identité respectivement avec *Marc* et *Lucas*.

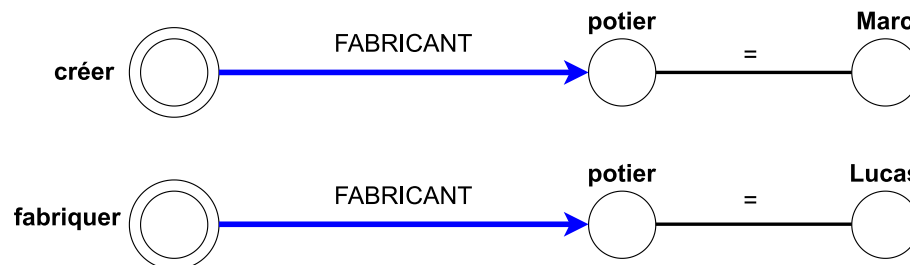


FIGURE 4.6 – Ces deux relations, qui sont contenues dans deux histoires distinctes, peuvent être alignées car il s’agit de deux relations de type FABRICANT existantes entre une action de fabrication et un humain.

En considérant toutes les relations qui sont contenues dans les deux histoires en (45) nous pouvons identifier d’autres relations qui peuvent être alignées : les deux histoires contiennent des relations de type OBJET FABRIQUÉ entre une action de fabrication (i.e. *fabriquer* et *créer*) et un résultat artefactuel (i.e. *pot* pour les deux histoires). Deuxièmement, les deux histoires

contiennent une relation d’ancrage temporel avec le moment, ou plus précisément la date, dans lequel l’action se déroule. Nous représentons en Figure 4.7 les relations des deux histoires en (45a) et (45b) que nous considérons comme étant alignables (i.e. arcs contigus) et les relations qui ne peuvent pas être alignées car elles sont présentes uniquement dans une histoire (i.e. arcs pointillés). L’alignement entre les relations sémantiques contenues dans les histoires est une étape préliminaire à l’abstraction de leur contenu et à l’obtention d’un *subset meaning bundle*, que nous détaillons en Section 4.9.

4.9 Subset Meaning bundles

Nous avons vu en Section 4.7 que les histoires construites sur des sous-ensembles de lexèmes d’une famille peuvent présenter des contenus sémantiques variés. Nous avons présenté les objets que nous utilisons dans les graphes sémantiques pour représenter les objets qu’elles contiennent (actions, activités, rôles, participants, propriétés). En outre, nous avons vu en Section 4.8 que les relations décrites par des histoires différentes peuvent être alignées en se basant sur des critères ontologiques et relationnels. Ces mêmes critères peuvent être utilisés pour aligner les relations sémantiques contenues dans toutes les histoires produites à partir d’un même sous-ensemble de lexèmes, comme $\{\text{POT.N}, \text{POTIER.N}\}$, pour identifier les relations qui sont récurrentes dans ces histoires. Par exemple, nous pouvons aligner toutes les relations de type FABRICANT impliquant un humain qui sont contenues dans les histoires produites à partir de $\{\text{POT.N}, \text{POTIER.N}\}$, ou encore toutes les relations de type OBJET FABRIQUÉ impliquant un artefact.

Une fois effectué l’alignement des relations au niveau de toutes les histoires produites à partir du même sous-ensemble, nous pouvons abstraire leur contenu. Pour faire cela, nous créons des relations abstraites en remplaçant les participants des histoires par une étiquette ontologique qui les regroupe et qui représente leur dénominateur commun. Par exemple, en Figure 4.8 pour effectuer l’abstraction de deux relations alignables contenues dans deux histoires distinctes, nous créons une relation abstraite de type FABRICANT en remplaçant les actions *fabriquer*

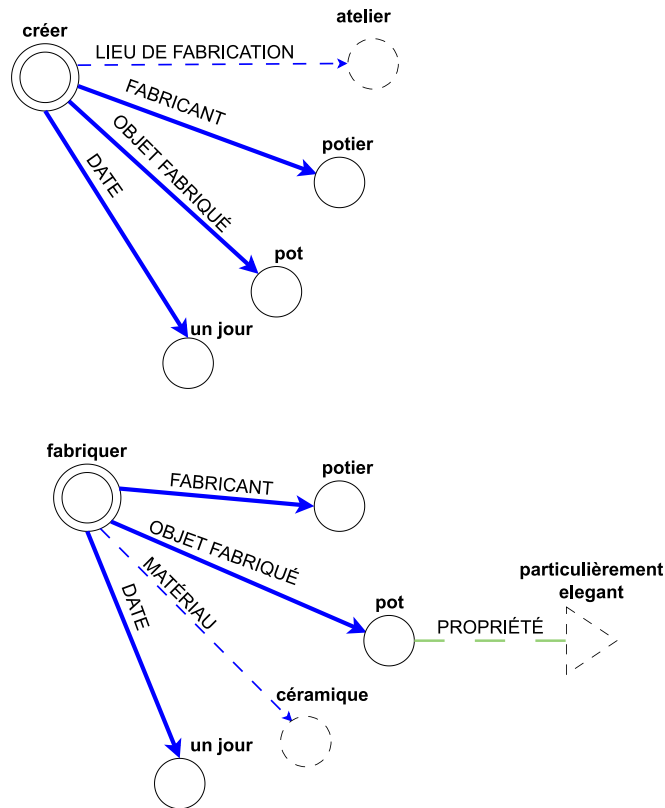


FIGURE 4.7 – En comparant les graphes des deux histoires en (45a) et (45b), nous considérons que trois relations peuvent être alignées sur une base ontologique et relationnelle : la relation de type FABRICANT que nous avons déjà illustrée en Figure 4.6, la relation de type OBJET FABRIQUÉ reliant les actions de création et leur résultat artefactuel (i.e. le pot) et la relation temporelle DATE reliant l’action au moment où elle se déroule. Les relations représentées par des lignes pointillées ne peuvent pas être alignées car elles ne sont pas présentes dans les deux histoires : la relation de type LIEU DE FABRICATION n’est pas présente dans l’histoire en (45b) et la relation entre l’action de fabrication et le matériau utilisé n’est pas présente dans l’histoire en (45a).

et *créer* par la catégorie ontologique ACTION DE FABRICATION et les participants humains des histoires par la catégorie PERSONNE¹⁶.

Nous considérons que si une relation sémantique est suffisamment fréquente dans les histoires produites à partir du même sous-ensemble de lexèmes d’une famille, nous pouvons l’abstraire et l’inscrire dans un premier type de *meaning bundle* que nous appelons *subset meaning bundle*

16. Comme pour les rôles, nous indiquons également les étiquettes ontologiques que nous utilisons en majuscule.

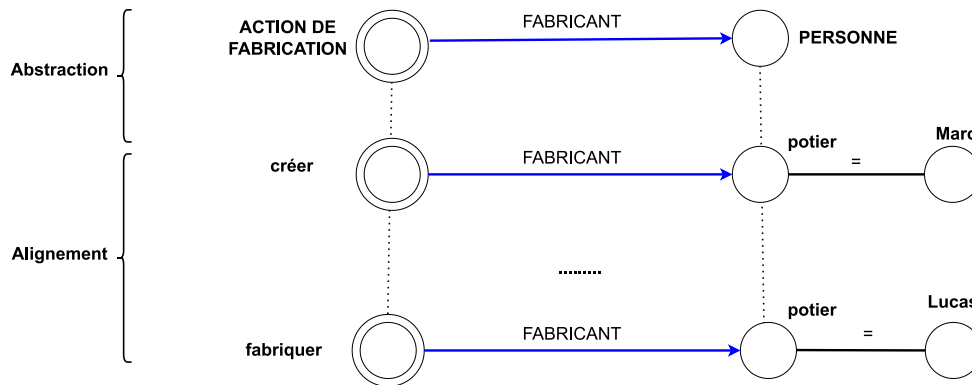


FIGURE 4.8 – Lorsque plusieurs relations sémantiques présentes dans différentes histoires peuvent être alignées, nous pouvons effectuer une abstraction de leur contenu. Les relations contenues dans les histoires construites à partir de $\{POT.N, POTIER.N\}$ de type FABRICANT entre un humain et une action de fabrication peuvent être abstraites en remplaçant les participants par leur catégorie ontologique qui représente leur dénominateur commun sur le plan ontologique. La ligne horizontale pointillée au centre de cette figure représente les autres relations que l'on peut aligner et dont le contenu peut être abstrait par la relation abstraite en haut de type FABRICANT entre une ACTION DE FABRICATION et une PERSONNE.

(SMB). Nous l'appelons de cette manière car il s'agit d'un *meaning bundle* contenant les participants et relations contenues régulièrement dans les histoires produites à partir du même sous-ensemble de lexèmes d'une famille (ex. $\{POT.N, POTIER.N\}$). Comme nous l'avons anticipé en Section 4.8, nous représentons tous les *meaning bundles*, et par conséquent les SMBs, au moyen de graphes.

Définition 9

Subset Meaning bundle

Un *subset meaning bundle* est un objet sémantique contenant les participants et les relations sémantiques qui coexistent régulièrement dans les histoires créées à partir d'un même sous-ensemble de lexèmes d'une famille dérivationnelle.

La question qui se pose maintenant est de déterminer dans combien d'histoires une relation entre deux participants doit apparaître pour pouvoir être considérée comme récurrente et être incluse dans un SMB. Pour cela, on peut envisager de fixer un seuil de fréquence qui nous permet de filtrer tout le contenu épisodique et de garder uniquement le contenu régulier des histoires.

Un exemple de contenu épisodique est fourni avec l’histoire en (46). L’événement dans lequel on s’attend qu’un pot et un potier interviennent est un événement de fabrication, alors que dans l’histoire en (46) le potier est la personne qui brise le pot. Dans l’histoire en (46) joue un rôle du “CASSEUR” et le pot joue le rôle de OBJET CASSÉ. Nous considérons que les relations décrites par l’histoire en (46) n’apparaîtront pas de manière assez fréquente dans les histoires construites sur {POT.N, POTIER.N} et seront filtrées par le seuil de fréquence.

(46) *"Le potier de mon village est très maladroit. L'autre jour, pendant qu'il travaillait, il a brisé un pot en mille morceaux."*

L’objectif est d’avoir un seuil de fréquence suffisamment haut pour ne garder que les relations récurrentes dans les histoires et qui apparaissent régulièrement ensemble. Cependant, comme la démarche que nous proposons est une démarche théorique, nous ne cherchons pas à quantifier ce seuil dans cette thèse.

Nous envisageons que parmi les histoires produites à partir du sous-ensemble {POT.N, POTIER.N} la grande majorité décrit des actions de fabrication qui impliquent un FABRICANT humain, un OBJET FABRIQUÉ artefactuel (pots, vases, bols), des MATÉRIAUX, un LIEU DE FABRICATION, des INSTRUMENTS DE FABRICATION de type artefactuel et des DÉCORATIONS. Nous avons pu voir tous ces participants intervenir dans les histoires que nous avons proposées jusqu’à ce point pour le couple {POT.N, POTIER.N}. En outre, nous considérons que, lorsqu’une histoire contient une action, cette action est délimitée d’un point de vue temporel et donc il existe une relation de type temporel (que nous avons étiqueté DATE dans les graphes présentés dans les Sections précédentes) entre l’action de fabrication et le moment où elle se déroule. Toutes ces relations sont représentées en Figure 4.9. Dans ce SMB nous n’incluons pas les propriétés associées aux participants, que nous avons vu apparaître dans quelques histoires, car nous considérons qu’elles ne sont pas centrales pour un scénario de fabrication et ne font pas partie de sa structure prototypique.

Comme on peut le voir en Figure 4.9, plusieurs participants inscrits dans un *meaning bundle*

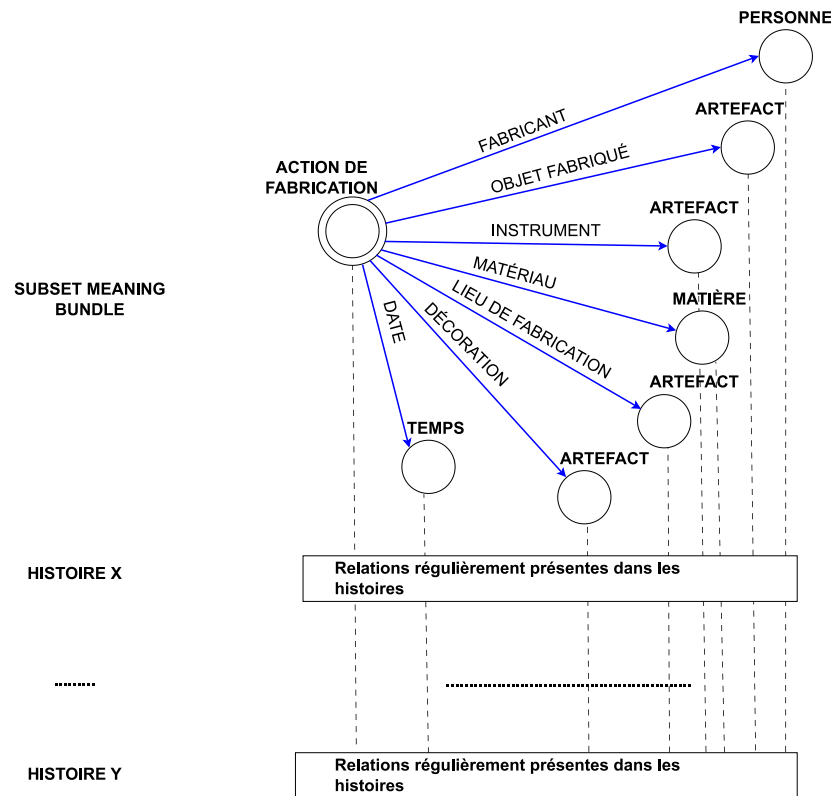


FIGURE 4.9 – SMB contenant les relations et les participants que nous considérons comme étant régulièrement présents dans les histoires créées à partir du sous-ensemble {POT.N, POTIER.N}. Le SMB est le résultat d’une abstraction (représentée par les lignes verticales pointillées) et d’un regroupement des relations suffisamment fréquentes dans les histoires produites à partir du même sous-ensemble et qui apparaissent régulièrement ensemble. Les lignes horizontales pointillées représentent respectivement l’ensemble d’histoires produites à partir du sous-ensemble {POT.N, POTIER.N} (ligne horizontale à gauche) et leur contenu régulier (ligne horizontale centrale).

peuvent partager des propriétés ontologiques : par exemple, l’OBJET FABRIQUÉ est un ARTEFACT (ex. pot, vase, bol) mais les INSTRUMENTS dont le FABRICANT se sert le sont aussi à priori (ex. tour). Il en va de même pour le LIEU DE FABRICATION (ex. atelier, usine). Le vrai élément discriminant est le rôle à l’intérieur de la situation : un atelier dans la plupart des cas n’est pas l’OBJET FABRIQUÉ dans une situation de fabrication, mais plutôt le LIEU DE FABRICATION (sauf dans des situations où on fabrique l’atelier lui-même) et un tour n’est pas l’OBJET FABRIQUÉ dans une situation de fabrication, mais plutôt l’INSTRUMENT DE

FABRICATION (sauf dans des situations où on fabrique le tour lui-même). Pour récapituler, dans un *meaning bundle* on peut avoir des participants qui ont la même nature ontologique mais qui jouent des rôles distincts dans la situation décrite par le MB.

En outre, nous précisons que les relations inscrites dans un *meaning bundle* (et donc dans un SMB) ne doivent pas être toujours présentes en même temps dans les histoires. Toutes les histoires n'ont pas en même temps le FABRICANT, l'OBJET FABRIQUÉ, le MATÉRIAU, le LIEU DE FABRICATION, l'INSTRUMENT et la DÉCORATION qui apparaissent systématiquement en même temps. La fréquence de ces relations au niveau de toutes les histoires produites à partir de {POT.N, POTIER.N} et le fait qu'elles apparaissent régulièrement l'une avec l'autre déterminent leur présence dans le SMB. En outre, ces relations partagent toutes un sommet : elles dépendent toutes d'une action de fabrication, qui est le pivot sémantique du réseau et qui attribue les rôles aux participants. Comme ces relations partagent toutes un même sommet et sont toutes suffisamment fréquentes, nous considérons qu'elles peuvent être regroupées dans le même graphe, à savoir celui qui est proposé en Figure 4.9.

En revanche, le SMB proposé en Figure 4.9 décrit uniquement des relations situées sur le plan action, alors que nous avons vu en Section 4.7 que les histoires créées à partir de {POT.N, POTIER.N} contiennent également des relations situées sur le plan activité et que le plan action et le plan activité peuvent coexister dans le cadre d'une histoire. Nous avons proposé en Section 4.7 de voir le réseau action et le réseau activité définis par Roché (2017b) comme deux réseaux qui au niveau sémantique ont la même structure et qui sont reliés par une relation d'instanciation orientée (i.e. l'action est une instance de l'activité). Sur cette base, nous proposons de compléter le SMB comme il est montré en Figure 4.10 pour rendre compte également des relations situées sur le plan activité qui sont contenues dans les histoires construites à partir de {POT.N, POTIER.N}.

Les deux composantes du réseau qui décrit le SMB en Figure 4.10 ont la même structure et impliquent le même type de participants. La seule différence est donnée par une relation de type temporel dans le réseau action qui n'est pas présente dans le réseau activité, car l'activité n'est pas ancrée d'un point de vue temporel. Cette situation va se représenter régulièrement

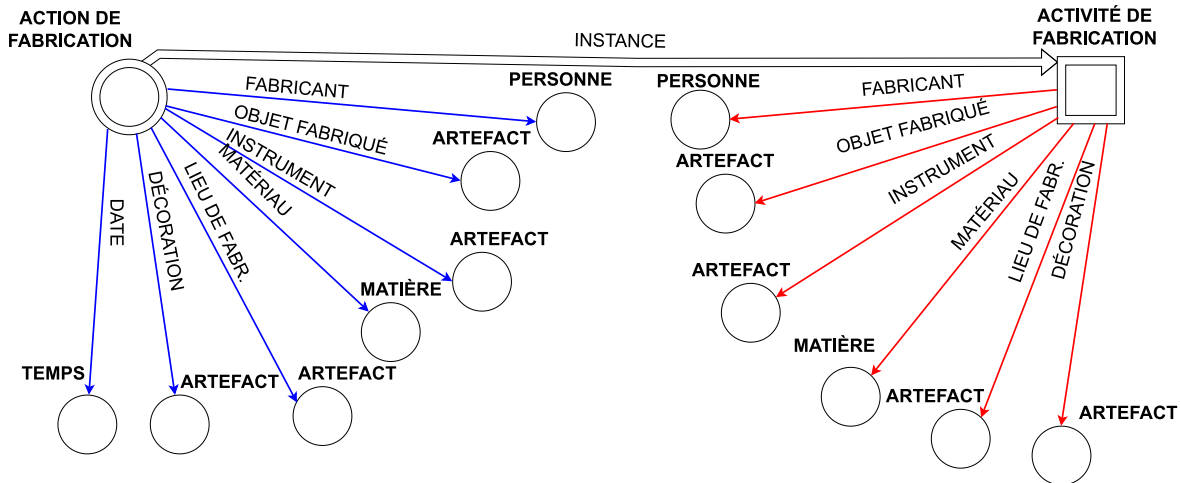


FIGURE 4.10 – SMB produit à partir du contenu sémantique des histoires construites sur le sous-ensemble $\{\text{POT.N, POTIER.N}\}$ comportant un volet action et un volet activité. Les deux réseaux qui composent ce SMB sont connectés via la relation d’instanciation qui relie le plan action au plan activité. Les deux plans partagent des sommets lorsque les participants situés sur le plan action sont les mêmes que ceux situés sur le plan activité.

pour d’autres *meaning bundles* comportant une composante activité et une composante action que nous allons présenter dans cette thèse. Pour rendre plus simples les graphes qui décrivent des *meaning bundles* comportant un volet action et un volet activité, à partir de la Section 4.10 nous les présentons en forme abrégée en représentant uniquement le volet activité et l’action qui l’instancie, sans répéter toutes les relations dans la composante action.

4.10 Traitement des autres sous-ensembles de la famille

La méthode que nous proposons prévoit de réaliser la même procédure de manière séparée pour tous les sous-ensembles de la famille $F1$ de départ. Pour illustrer le résultat de cette étape, nous considérons les sous-ensembles proposés en (47). Ces sous-ensembles introduisent des cas de figure variés.

- (47) a. $\{\text{POT.N, POTERIE.N}\}$
 b. $\{\text{POT.N, REMPOTER.V, REMPOTEUR.N}\}$

c. {POT.N, POTIER.N, POTERIE.N, REMPOTER.V, REMPOTEUR.N}

{POT.N, POTERIE.N} Nous considérons que les histoires produites à partir de {POT.N, POTERIE.N} contiennent de manière régulière des relations et des participants que l'on peut considérer comme étant équivalents à ceux que nous avons identifiés pour le SMB produit à partir des histoires construites sur {POT.N, POTIER.N}, qui est montré en Figure 4.10. Un exemple d'histoire produite à partir de {POT.N, POTERIE.N} est proposé en (48) : l'histoire décrit une activité de fabrication de pots avec un FABRICANT qui s'y consacre, la céramique qui intervient en tant que MATÉRIAU et des DÉCORATIONS (i.e. les formes élégantes et les motifs créatifs) qui sont réalisées sur le pot.

(48) *Lucien était un talentueux artisan qui se consacrait à la poterie. Il travaillait avec passion dans son atelier, façonnant des pots en céramique aux formes élégantes et aux motifs créatifs. Chaque création était unique, témoignant de son savoir-faire et de sa créativité.*

En considérant l'histoire en (48) comme étant illustrative de la grande majorité des histoires que l'on peut produire à partir de {POT.N, POTERIE.N}, nous envisageons que le SMB que l'on peut abstraire à partir de ce sous-ensemble est équivalent à celui que nous avons proposé en Figure 4.10 pour {POT.N, POTIER.N}. Il en va de même pour le traitement d'autres sous-ensembles comme {POTIER.N, POTERIE.N} et {POT.N, POTIER.N, POTERIE.N}.

{POT.N, REMPOTER.V, REMPOTEUR.N} Les histoires créées à partir de ce sous-ensemble traitent d'actions ou d'activités de rempotage, à savoir de changements de localisation affectant une plante. Nous proposons un exemple d'histoire illustratif en (49). Comme on peut le voir, cette histoire contient également REMPOTAGE, que nous n'avons pas pris en compte dans la famille initiale et ni dans le prompt donné à *ChatGPT*, car il dénote le même processus que REMPOTER.

(49) *Vendredi le rempoteur Robin a décidé de rempoter ses cactus dans des pots plus grands. Avec précaution, il a effectué le rempotage, ajoutant de la terre fraîche pour favoriser la*

croissance. Le rempotage a été un succès, les plantes ont rapidement retrouvé leur vigueur dans leurs nouveaux environnements.

Nous proposons une représentation simplifiée du contenu sémantique de l’histoire en (49) en Figure 4.11 en nous concentrant sur l’action de déplacement de la plante. Dans ce cas, l’histoire est structurée par l’action de rempoter, à savoir le déplacement de la plante dans le nouveau pot. Pour relier l’action de rempoter à ses participants, nous proposons d’utiliser les rôles sémantiques suivants :

- AGENT DU DÉPLACEMENT : il s’agit d’une personne qui déplace de manière volontaire un objet (i.e. *le rempoteur Robin*);
- ENTITÉ DÉPLACÉE : entité qui subit un déplacement (i.e. *les cactus*);
- DESTINATION : lieu ou récipient d’arrivée d’un déplacement (i.e. *les pots*);
- MOYEN : entité qui est utilisée pour réaliser une action et qui est modifiée par le procédé dans lequel elle intervient (Fradin, 2021, p.161) (i.e. *la terre fraîche*).

Nous considérons que la plupart des histoires produites à partir du sous-ensemble {POT.N, REMPOTER.V, REMPOTEUR.N} décrivent des actions ou des activités de rempotage de la même nature que celle qui est décrite par l’histoire en (49). Sur cette base, nous considérons que le SMB produit à partir de {POT.N, REMPOTER.V, REMPOTEUR.N} contient les relations qui sont proposées en Figure 4.12. Comme nous l’avons fait pour l’action de fabrication précédemment en Section 4.7, nous attribuons des étiquettes ontologiques aux participants du SMB pour abstraire le contenu des histoires : dans ce cas nous considérons que dans la grande majorité des histoires construites à partir de ce sous-ensemble l’AGENT DU DÉPLACEMENT rentre dans la catégorie PERSONNE, l’ENTITÉ DÉPLACÉE rentre dans la catégorie PLANTE, la DESTINATION rentre dans la catégorie POT et on aura du TERREAU qui est utilisé comme MOYEN pour rempoter la plante. En outre, nous considérons que l’action de rempotage peut être vue comme une instance d’une activité de rempotage, que l’on peut inscrire dans une activité de jardinage. Le SMB en Figure 4.12 contient donc également deux volets (action et activité) qui ont la même structure et qui sont reliés par une relation d’instanciation.

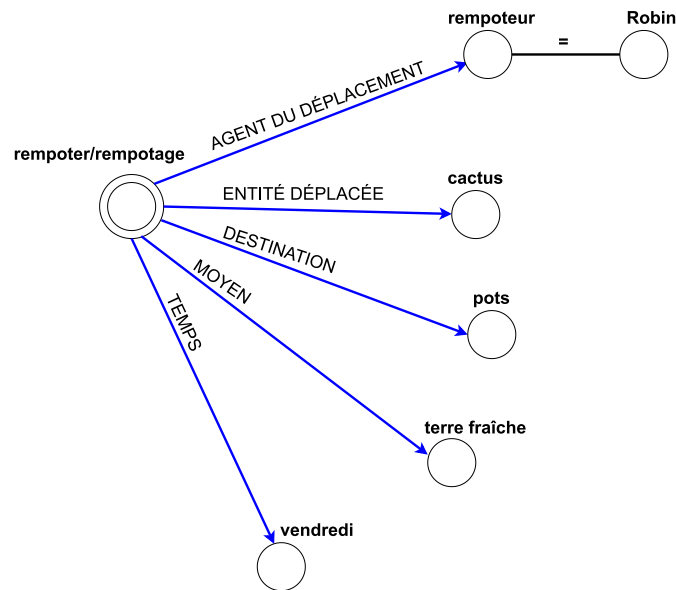


FIGURE 4.11 – Représentation simplifiée du contenu sémantique de l'histoire (49). L'histoire décrit un déplacement : pour cette raison nous attribuons le rôle de AGENT DU DÉPLACEMENT au rempoteur et de ENTITÉ DÉPLACÉE aux cactus. Dans cette histoire, les pots sont la DESTINATION du déplacement. Concernant la terre, tout en étant elle aussi déplacée dans le processus, nous considérons que sa fonction principale est plutôt celle de MOYEN qui est nécessaire à la croissance de la plante.

En outre, nous considérons que des SMBs équivalents à celui en Figure 4.12 peuvent être obtenus à partir de couples comme $\{\text{REMPOTER.V, REMPOTEUR.N}\}$, $\{\text{POT.N, REMPOTEUR.N}\}$ et $\{\text{POT.N, REMPOTER.V}\}$.

$\{\text{POT.N, POTIER.N, POTERIE.N, REMPOTER.V, REMPOTEUR.N}\}$ Nous envisageons que les histoires produites à partir de ce sous-ensemble décrivent principalement des actions et activités de fabrication de poterie, des activités de repotage dans un cadre de jardinage ou des combinaisons particulières des deux. Un exemple qui va dans ce sens est l'histoire en (50). Cette histoire décrit à la fois une activité régulière de fabrication de pots et une action de repotage. Pour les distinguer dans l'histoire en (50), nous marquons en rouge les parties qui font référence à des participants de l'activité de fabrication des pots et en bleu les parties qui font référence aux participants de l'action de repotage. Ces deux composantes de l'histoire sont connexes,

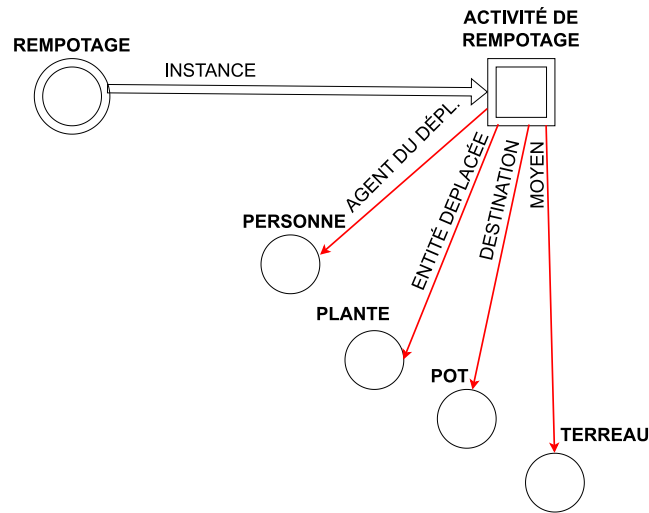


FIGURE 4.12 – SMB obtenu à partir des histoires construites sur le sous-ensemble {POT.N, REMPOTER.V, REMPOTAGE.N}. Ce SMB décrit une action de déplacement d’une plante dans un pot qui instancie une activité de rempotage et, comme le SMB proposé précédemment en Figure 4.10, contient des relations situées sur le plan action et des relations situées sur le plan activité.

car les nouveaux pots dont l’histoire parle sont impliqués à la fois dans l’activité de fabrication et dans l’action de rempotage que le potier réalise : pour cette raison, *nouveaux pots* est marqué avec les deux couleurs en (50). Il en va de même pour le potier, qui est le FABRICANT par rapport à l’activité de fabrication des pots et l’AGENT DU DÉPLACEMENT par rapport au rempotage. En outre, *potier* est relié par une relation d’identité avec *remporteur*, car il s’agit de la même personne.

- (50) *Un potier talentueux nommé Pierre consacrait sa vie à la poterie. Chaque jour, il façonnait des pots de toutes formes et tailles et les décorait avec des motifs uniques. Un jour, Pierre s’improvisa remporteur et décida de rempoter ses plantes dans de nouveaux pots qu’il avait créés lui-même, réalisant ainsi un rempotage soigneux. Avec précision, il transféra chaque plante dans sa nouvelle demeure en terre cuite, savourant l’union parfaite entre son art de la poterie et sa passion pour le jardinage.*

Comme toutes les histoires sémantiquement cohérentes, l’histoire en (50) est formalisable sous forme de graphe connexe, comme nous le montrons en Figure 4.13. Le point de connexion

est le potier Pierre, qui est impliqué dans les deux réseaux. En outre, les pots qui dans l'histoire en (50) sont utilisés pour le rempotage sont un sous-ensemble des pots que le potier a produit dans le cadre de son activité. Nous représentons cette relation au moyen d'un arc pointillé orienté typé avec le symbole \subset , qui équivaut à 'sous-ensemble'.

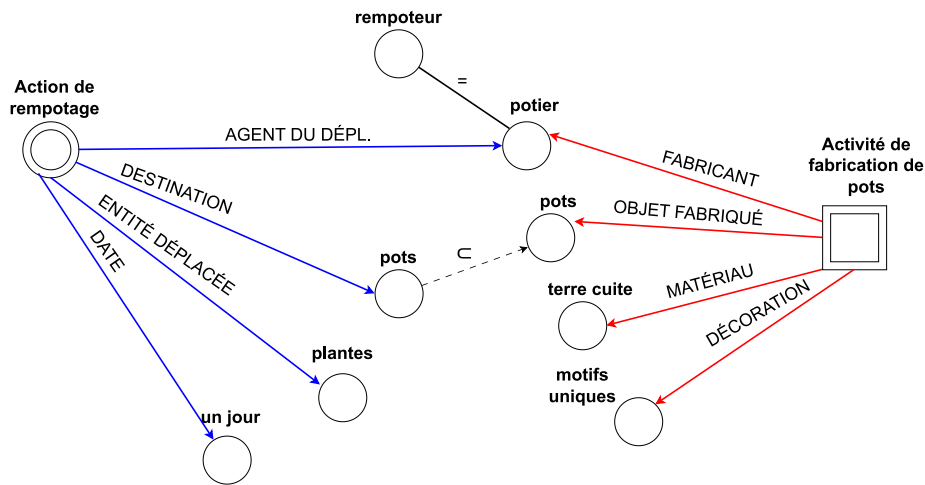


FIGURE 4.13 – Représentation simplifiée du contenu sémantique de l'histoire en (50). Cette histoire décrit une activité de fabrication de pots et une action de rempotage réalisée par le même potier en utilisant un sous-ensemble des pots qu'il a fabriqués. Les deux composantes partagent donc le sommet référant au potier, qui joue en même temps deux rôles distincts dans l'histoire.

En considérant l'histoire proposée en (50) comme étant illustrative de la majorité des histoires produites à partir de ce sous-ensemble, nous envisageons que l'alignement et l'abstraction des relations contenues dans les histoires construites sur $\{\text{POT.N, POTIER.N, POTERIE.N, REMPOTER.V, REMPOTEUR.N}\}$ produisent un SMB connexe qui contient à la fois une action ou une activité de fabrication des pots et une action ou une activité de rempotage. Nous envisageons que la connexion entre la composante qui décrit la fabrication d'un pot et la composante qui décrit le rempotage se fera toujours aux mêmes endroits : **(a)** le pot qu'on fabrique dans l'histoire est le même qui est utilisé pour le rempotage et/ou **(b)** la personne qui a fabriqué le pot est la même qui réalise le rempotage. Une schématisation du SMB obtenu à partir du sous-ensemble $\{\text{POT.N, POTIER.N, POTERIE.N, REMPOTER.V, REMPOTEUR.N}\}$ est proposée en Figure 4.14.

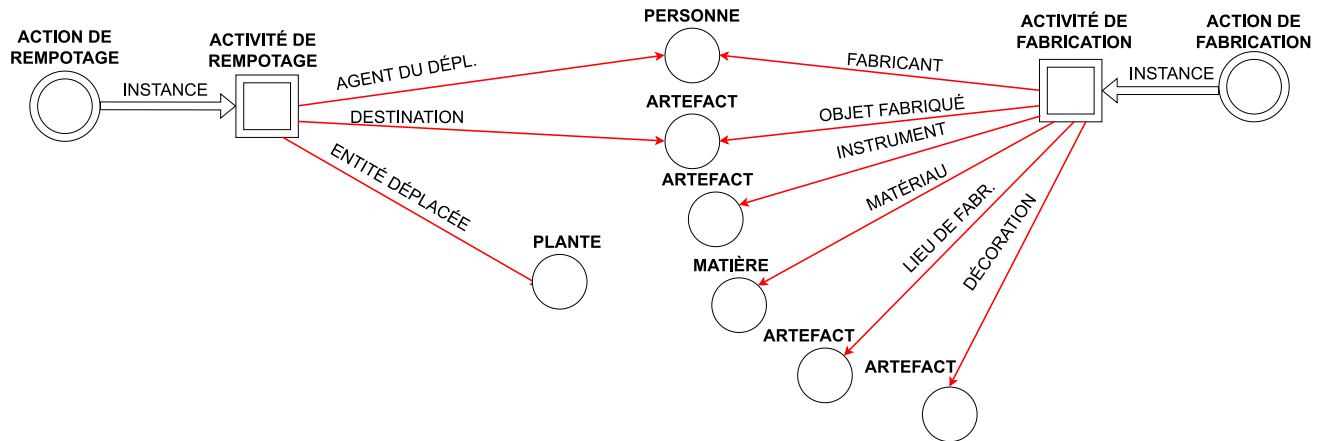


FIGURE 4.14 – SMB produit à partir du contenu des histoires construites sur le sous-ensemble $\{\text{POT.N, POTIER.N, POTERIE.N, REMPOTER.V, REMPOTAGE.N}\}$. Ce SMB comporte une composante décrivant une action/activité de rempotage et une action/activité de fabrication.

En revanche, la coexistence régulière d'une action/activité de fabrication d'un pot et d'une action/activité de rempotage qui est décrite par le SMB en Figure 4.14 est forcée par le sous-ensemble que nous avons pris en compte et se vérifiera uniquement dans d'autres sous-ensembles similaires, comme $\{\text{POT.N, POTIER.N, REMPOTER.V}\}$ ou $\{\text{POT.N, POTIER.N, REMPOTER.V, REMPOTEUR.N}\}$. Des SMBs comme celui en Figure 4.14 vont à l'encontre de notre objectif d'identifier des paquets de relations interprédictibles : nous gérons ce problème en montant en termes d'abstraction et en considérant les participants et les relations récurrentes au niveau de toutes les histoires produites à partir de tous les sous-ensembles de la famille dont nous sommes partis. Cela nous amène aux *family meaning bundles* (FMB).

4.11 Meaning bundles au niveau d'une famille

Nous avons des SMBs pour chaque sous-ensemble contenu dans la famille dérivationnelle de POT. Comment obtenir des *meaning bundles* au niveau de toute la famille ? Pour faire cela, nous proposons d'aligner les SMBs que nous avons obtenus dans l'étape précédente lorsqu'ils contiennent des relations sémantiques alignables et lorsque ces relations sont suffisamment fréquentes au niveau de toutes les histoires construites sur tous les sous-ensembles de lexèmes de la

famille. Nous appelons les *meaning bundles* obtenus au moyen de cet alignement *family meaning bundles* (FMB).

Définition 10**Family Meaning bundle**

Un *family meaning bundle* (FMB) est un objet sémantique contenant un ensemble de participants et de relations qui interviennent régulièrement ensemble dans toutes les histoires produites à partir de tous les sous-ensembles d'une famille dérivationnelle.

Pour effectuer l'alignement entre SMBs, nous nous basons sur les mêmes critères relationnels et ontologiques que ceux que nous avons adoptés en Section 4.8. En outre, nous utilisons à nouveau un seuil de fréquence pour ne garder dans le même FMB que les participants et les relations que dans toutes les histoires produites à partir de tous les sous-ensembles **(a)** sont suffisamment fréquentes **(b)** apparaissent ensemble de manière suffisamment fréquente.

Nous rappelons que le traitement de tous les sous-ensembles de la famille de POT dont nous sommes partis produit des SMBs qui décrivent une action/activité de fabrication, une action/activité de repotage, ou les deux en même temps, comme dans le cas du sous-ensemble {POT.N, POTIER.N, POTERIE.N, REMPOTER.V, REMPOTEUR.N} que nous venons de traiter ou dans d'autres sous-ensembles comme {POT.N, POTIER.N, REMPOTER.V}. En revanche, nous considérons que les seuls SMBs dans lesquels un scénario de fabrication et de repotage interviennent ensemble sont ceux où cette coexistence est forcée par le sous-ensemble que nous prenons en compte (voir Section 4.10). Au contraire, au niveau de toutes les histoires produites à partir de tous les sous-ensembles de la famille de POT, le scénario de fabrication et celui qui décrit un repotage n'apparaissent pas assez fréquemment ensemble pour pouvoir les considérer comme un tout et les inscrire dans un même FMB connexe. Les relations qui permettraient aux deux scénarios de coexister (ex. un pot qui a été fabriqué est le même qui est utilisé pour le repotage) ne seront pas fréquentes au niveau de toutes les histoires, mais le seront uniquement au niveau des histoires créées à partir des sous-ensembles qui induisent cette coexistence. L'usage d'un seuil de fréquence dans cette phase permet de filtrer ces relations.

D'autre part, nous considérons que dans la plupart des histoires produites à partir de tous

les sous-ensembles de la famille de POT un FABRICANT, un OBJET FABRIQUÉ, des MATÉRIEAUX, un LIEU DE FABRICATION, des INSTRUMENTS et des DÉCORATIONS coexistent régulièrement dans la même histoire. En outre, au niveau de toutes les histoires créées à partir de tous les sous-ensembles, si nous avons un rempotage dans une histoire, nous avons régulièrement un AGENT DU DÉPLACEMENT, une ENTITÉ DÉPLACÉE et une DESTINATION.

Cela signifie qu’au niveau de toutes les histoires produites à partir de tous les sous-ensembles de la famille de POT on peut produire deux FMBs non connexes : un qui décrit une action/activité de fabrication d’un pot et l’autre qui décrit une action/activité de rempotage. Ces deux FMBs non connexes sont proposés en Figure 4.15 : chacun de ces FMBs contient des participants qui sont interprédictibles. Les deux FMBs en Figure 4.15 sont pertinents pour deux facettes distinctes du sens de POT : sa nature d’OBJET FABRIQUÉ est pertinente pour l’action/l’activité de fabrication, alors que sa fonction de récipient est pertinente pour l’action/activité de rempotage, dans laquelle le pot joue le rôle de la DESTINATION.

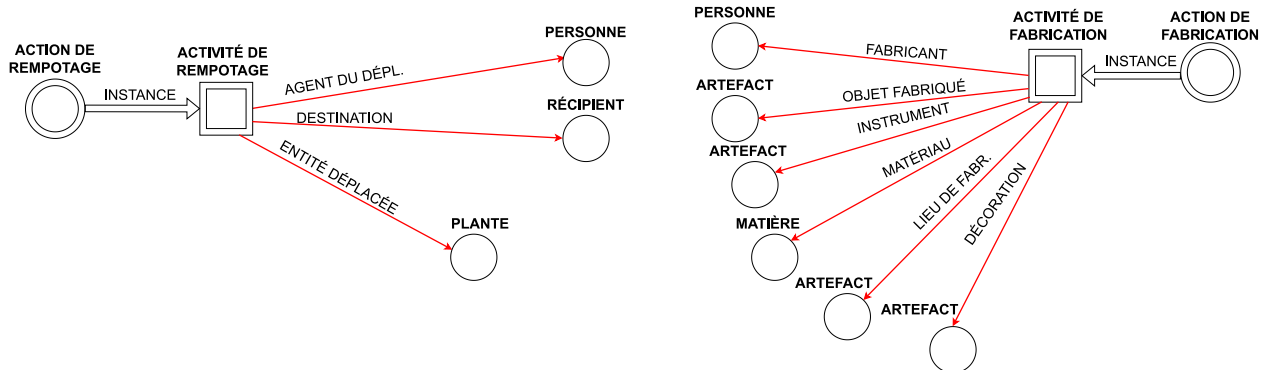


FIGURE 4.15 – FMBs non connexes obtenus à partir de la famille de POT d’où nous sommes partis. Le passage de SMB à FMB a permis de filtrer les relations qui connectaient le réseau décrivant un scénario de fabrication et celui qui décrit le rempotage. Cette figure présente donc deux FMBs séparés, contrairement à la Figure 4.14, qui représente un SMB qui réunit les deux situations.

La fonction du passage des SMBs aux FMBs est de filtrer toutes les coexistences accidentelles de participants qui sont induites par l’étape précédente, où nous avons pris en compte tous les sous-ensembles possibles de la famille de départ, y compris des sous-ensembles qui mélangent des scénarios distincts et qui ne sont interprédictibles (ex. la fabrication d’un pot et le rempotage).

Lorsque les lexèmes contenus dans une famille de départ s'inscrivent dans deux ou plusieurs scénarios distincts, à partir de cette famille on peut produire deux ou plusieurs FMBs. En prenant en compte la famille de POT dont nous sommes partis, nous en avons obtenu deux, à savoir ceux qui sont proposés en Figure 4.15.

4.12 FMBs obtenus à partir d'autres familles et LMBs

Nous disposons maintenant de deux FMBs non connexes obtenus à partir de la famille dérivationnelle de POT. Nous rappelons l'objectif que nous nous sommes donnés, à savoir identifier des paradigmes dans le lexique construit français. Les paradigmes que nous voulons identifier contiennent des relations récurrentes dans le lexique : cela signifie que nous devons identifier dans le lexique construit d'autres familles qui contiennent les mêmes relations que celles qui sont inscrites dans les deux FMBs que nous avons identifiés dans la famille de POT. Comme nous ne savons pas quelles familles contiennent les relations que nous cherchons, nous utilisons la même approche de force brute citée précédemment et nous prenons en compte toutes les familles du lexique en construisant des histoires sur tous leurs sous-ensembles pour révéler les relations que ces familles contiennent. De cette manière, nous pouvons obtenir des FMBs qui peuvent être alignés avec un des deux FMBs que nous venons de produire en Section 4.11.

En supposant que la même méthode que nous avons illustrée jusqu'ici pour la famille de POT soit faite pour toutes les familles du lexique construit, à ce stade nous disposons d'un ou plusieurs FMBs obtenus à partir de chaque famille du lexique. Nous proposons d'aligner les FMBs qui contiennent les mêmes relations en adoptant les mêmes critères ontologiques et relationnels utilisés pour les alignements précédents. Cette étape nous permet d'obtenir des *lexical meaning bundles* (LMB). Nous les appelons de cette manière car ils contiennent des régularités sémantiques au niveau de toutes les histoires produites à partir de toutes les familles présentes dans le lexique construit.

Définition 11**Lexical meaning bundle**

Un *lexical meaning bundle* (LMB) est un objet sémantique contenant un ensemble de participants et de relations sémantiques qui apparaissent ensemble régulièrement au niveau de toutes les histoires produites à partir de toutes les familles du lexique construit.

Pour illustrer le déroulement de cette étape, nous considérons les familles partielles qui sont proposées en (51) et nous fournissons des exemples d’histoires que l’on peut produire à partir de ces familles partielles. Sur la base des propriétés sémantiques des lexèmes contenus en (51), nous considérons qu’à partir de ces familles nous pouvons produire des FMBs qui contiennent les mêmes relations que celles que nous avons identifiées dans les FMBs obtenus à partir de la famille de POT. Nous justifions cette considération pour chaque famille partielle en (51) dans la suite de cette section. Pour la famille partielle de BOTTE, contrairement à la Section 4.2, nous distinguons le sens de ‘activité de fabrication de bottes’ (BOTTERIE.N₁) et ‘atelier de fabrication de bottes’ (BOTTERIE.N₂). Il en va de même pour FAÏENCE, qui dénote à la fois le matériau (FAÏENCE₁) et un objet construit en faïence (FAÏENCE₂), et pour FAÏENCERIE, qui dénote des objets en faïence (FAÏENCERIE₁), l’activité de fabrications d’objets en faïence (FAÏENCERIE₂) et un lieu où l’on fabrique des objets en faïence (FAÏENCERIE₃). En (51b), le sens de TOURNER qui est considéré est ‘façonner des pièces à l’aide d’un tour’. En (51d), le sens de EMBOUTEILLER qui est considéré est ‘mettre en bouteille’.

- (51) a. $F2 = \{\text{BOTTE.N, BOTTIER.N, BOTTERIE.N}_1 \text{ (activité), BOTTERIE.N}_2 \text{ (lieu)}\}$
 b. $F3 = \{\text{TOUR.N, TOURNEUR.N, TOURNER.V}\}$
 c. $F4 = \{\text{FAÏENCE.N}_1 \text{ (matériau), FAÏENCE.N}_2 \text{ (artefact), FAÏENCIER.N, FAÏENCERIE.N}_1 \text{ (artefact), FAÏENCERIE.N}_2 \text{ (activité), FAÏENCERIE.N}_3 \text{ (lieu)}\}$
 d. $F5 = \{\text{BOUEILLE.N, EMBOUTEILLER.V, EMBOUTEILLAGE.N, EMBOUTEILLEUR.N}\}$

{BOTTE.N, BOTTIER.N, BOTTERIE.N₁, BOTTERIE.N₂} Nous considérons qu’un grand nombre d’histoires produites à partir de la famille de BOTTE, dont le sous-ensemble {BOTTE.N, BOTTIER.N, BOTTERIE.N₁, BOTTERIE.N₂} fait partie, décrivent une activité de fabrication de bottes

ou de chaussures similaires ou une action qui l'instancie. Un exemple illustratif d'histoire produit à partir de ces lexèmes est fourni en (52).

(52) *Dans une petite botterie de Paris, un bottier talentueux travaillait avec passion. Cette nuit-là, il créa une botte en cuir d'une élégance sans précédent, ornée de motifs délicats.*

Nous considérons donc qu'à partir de la famille de BOTTE, qui contient les lexèmes en (51a), nous pouvons produire un FMB comme celui en Figure 4.16. Ce FMB est structuré par les mêmes relations que celui qui décrit la fabrication d'un pot ou un artefact similaire : il contient un FABRICANT humain, l'OBJET FABRIQUÉ est une chaussure (botte, bottine, etc.), il prévoit la présence d'INSTRUMENTS artefactuels (ex. outils servant au lissage, à guillocher le rebord de la semelle, etc.), il prévoit des MATÉRIAUX (ex. cuire), un LIEU DE FABRICATION artefactuel (ex. botterie, atelier) et des DÉCORATIONS qui sont réalisées sur la chaussure. Sur cette base, nous considérons que ce FMB peut être aligné avec celui qui décrit une activité de fabrication de pots que nous avons présentée en Section 4.11.

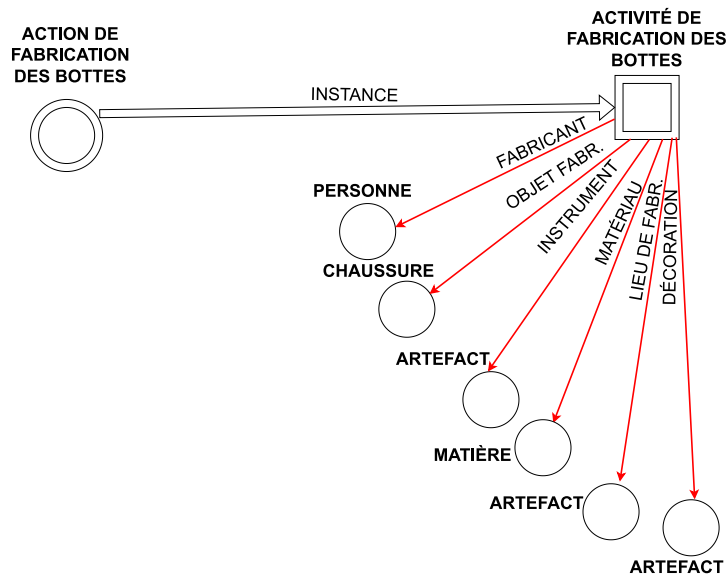


FIGURE 4.16 – FMB produit à partir de la famille de BOTTE structuré par une activité de fabrication de bottes ou d'autres chaussures.

{TOUR.N, TURNER.V, TOURNEUR.N} Nous considérons qu'à partir de la famille de TOUR, dont ce sous-ensemble fait partie, on peut produire un FMB qui décrit une activité ou une action de fabrication d'une pièce au tour. Un exemple d'histoire illustratif est proposé en (53).

- (53) *Au cœur d'un atelier poussiéreux, Luc, un tourneur habile, manipulait avec dextérité son tour. Chaque jour, il choisissait un bloc de bois qu'il aimait tourner pour le transformer en une œuvre d'art.*

Sur cette base, nous considérons qu'un des FMBs que l'on peut produire à partir de la famille de TOUR contient des relations de même nature que celles que nous avons identifiées dans les FMBs créés à partir des familles de POT et de BOTTE qui décrivent des actions/activités de fabrication. On a un FABRICANT (le tourneur), un OBJET FABRIQUÉ (les pièces que l'on fabrique au tour), l'INSTRUMENT (le tour lui-même), le MATÉRIAU (bois, métal ou autre), un LIEU DE FABRICATION (un atelier ou une salle dédiée). Comme le tour est utilisé pour fabriquer principalement des pièces mécaniques, il est moins probable que les histoires construites sur {TOUR.N, TURNER.V, TOURNEUR.N} impliquent régulièrement des DÉCORATIONS. Pour cette raison, nous les excluons du FMB en Figure 4.17.

D'autre part, le sous-ensemble {TOUR.N, TURNER.V, TOURNEUR.N} présente une configuration différente par rapport aux familles de POT et de BOTTE que nous avons considérées précédemment : dans le FMB que nous venons de proposer, le tour ne joue pas le rôle de OBJET FABRIQUÉ, mais plutôt d'INSTRUMENT. Le rôle prototypique joué par l'élément initial de la famille de TOUR dans une situation de fabrication (i.e. INSTRUMENT) est différent du rôle joué par POT et BOTTE (i.e. OBJET FABRIQUÉ), mais cela n'empêche pas de pouvoir créer des histoires de fabrication à partir de ces trois familles.

{FAÏENCE.N₁, FAÏENCE.N₂, FAÏENCIER.N, FAÏENCERIE.N₁, FAÏENCERIE.N₂, FAÏENCERIE.N₃} Nous considérons qu'à partir de la famille de FAÏENCE, dont ce sous-ensemble fait partie, on peut produire également de manière régulière des histoires qui décrivent la fabrication de faïence (vases, plats, carafes). Un exemple d'histoire est proposé en (54).

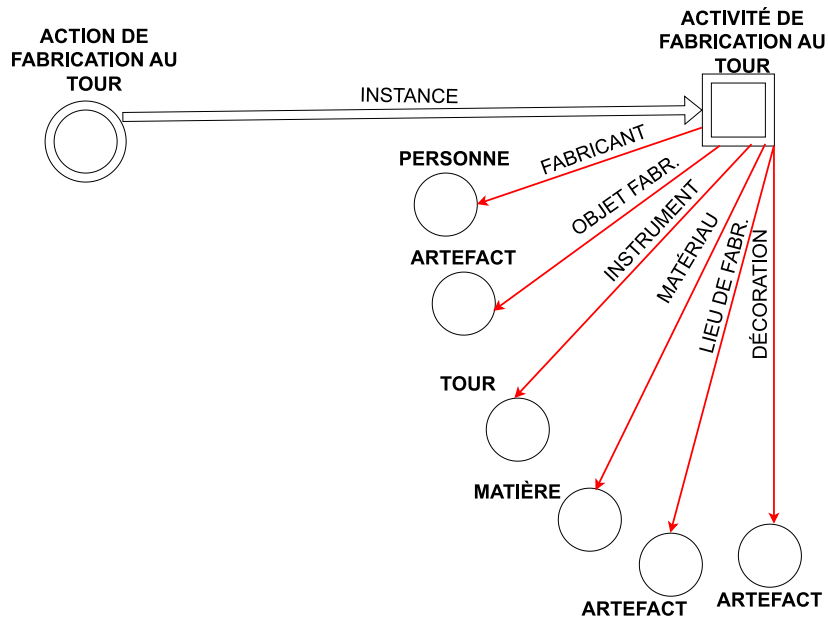


FIGURE 4.17 – FMB obtenu à partir de la famille de TOUR qui décrit l'activité de fabrication d'un artefact au tour et l'action qui l'instancie.

- (54) *Le talentueux faïencier Pierre dirigeait une faïencerie renommée. Chaque jour, il produisait des œuvres d'art en faïence. Les visiteurs venaient de loin pour admirer les créations uniques de sa faïencerie, qui étaient exposées dans un atelier.*

Sur cette base, nous considérons qu'à partir de la famille de FAÏENCE nous pouvons produire également un FMB qui décrit une activité de fabrication d'un artefact. Ce FMB est proposé en Figure 4.18 : nous considérons que les histoires produites à partir de cette famille contiennent régulièrement un FABRICANT (le faïencier), des OBJETS FABRIQUÉS (des faïences), des INSTRUMENTS (ex. un four, outils pour le modelage), un MATÉRIAU (la faïence), un LIEU DE FABRICATION (atelier, usine ou autre) et des DÉCORATIONS qui peuvent être réalisées sur l'objet (dessins, gravures).

Ce qui distingue la famille partielle de FAÏENCE de celles de POT, BOTTE et TOUR est le fait que l'élément initial de la famille, (FAÏENCE), joue le rôle du MATÉRIAU dans la fabrication, car le sens primaire de FAÏENCE est le sens 'matière'. Sur le plan sémantique, d'autre part, cela n'empêche pas de produire des histoires qui décrivent des actions ou des activités de fabrication

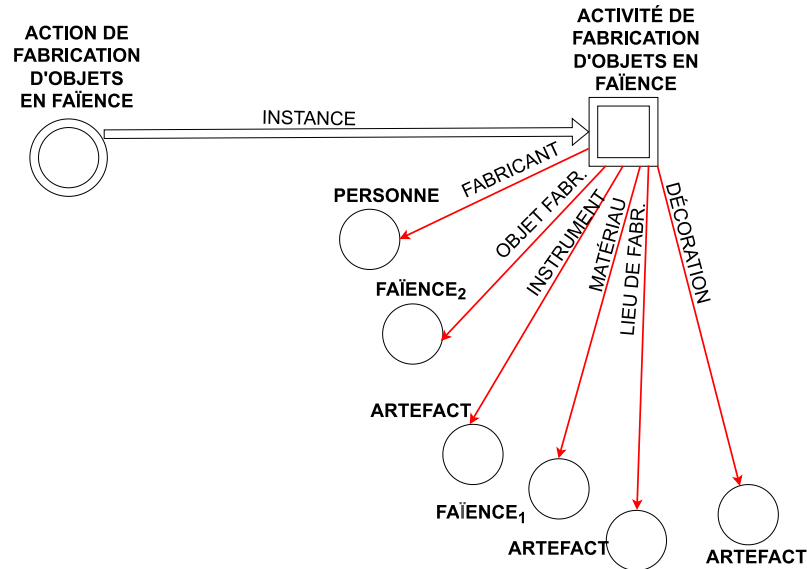


FIGURE 4.18 – FMB produit à partir de la famille de FAÏENCE qui décrit l’activité de fabrication d’objets en FAÏENCE.

et d’obtenir un FMB comme celui en Figure 4.18.

{BOUTEILLE.N, EMBOUTEILLER.V, EMBOUTEILLEUSE.N, EMBOUTEILLEUR.N} Pour sa part, la famille partielle de BOUTEILLE en (51d) contient un nom de récipient, l’action de mettre une substance dans ce récipient, le nom de l’instrument utilisé pour effectuer cette action et la personne chargée de réaliser l’action. La propriété sémantique qui est pertinente pour la structuration de cette famille n’est pas la nature artefactuelle de la bouteille, mais plutôt sa fonction de récipient. Sur la base des propriétés ontologiques et relationnelles des lexèmes des familles en (51) nous pouvons envisager que les histoires produites à partir de cette famille décrivent principalement des activités ou des actions de mise en bouteille de substances effectuées par des humains ou par des machines, par exemple à l’aide d’une embouteilleuse. Un exemple d’histoire de ce type est proposé en (55).

- (55) *L’embouteilleuse moderne était une machine impressionnante, capable de remplir des milliers de bouteilles par heure. Les embouteilleurs de la brasserie travaillaient sans relâche pour embouteiller la bière fraîchement brassée, assurant ainsi sa distribution à*

grande échelle. Chaque bouteille, soigneusement scellée, était prête à être savourée par les amateurs de bière du monde entier.

Sur cette base, nous considérons que le FMB produit à partir de la famille de BOUTEILLE en (51d) contient au moins les relations représentées en Figure 4.19. La différence par rapport au FMB décrivant le rempotage que nous avons identifié précédemment est le fait que ce FMB contient également l'INSTRUMENT utilisé pour réaliser l'action de déplacement (une embouteilleuse), alors que normalement le rempotage se fait manuellement. Nous considérons que dans ces histoires la personne qui effectue l'embouteillage est l'AGENT DU DÉPLACEMENT et la bouteille joue le rôle de la DESTINATION. En dernier lieu, en Figure 4.19 nous considérons que l'ENTITÉ DÉPLACÉE est généralement une SUBSTANCE, à savoir le contenu prototypique des bouteilles (55).

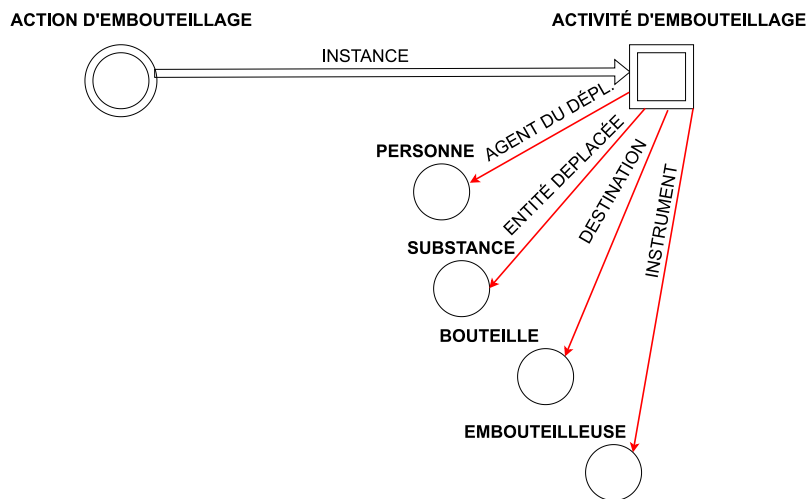


FIGURE 4.19 – FMB produit à partir de la famille de BOUTEILLE en (51d). La différence par rapport au FMB décrivant le rempotage est la présence d'une machine qui peut accompagner ou remplacer l'humain dans l'activité.

Un autre problème se pose à ce stade : quel niveau de généralisation choisir pour l'alignement des FMBs ? Jusqu'à quel point on accepte de généraliser les relations qui sont contenues dans les familles ? Nous considérons que le bon niveau de généralisation est révélé par le contenu des familles du lexique et des histoires que l'on peut produire à partir de ces familles : ce n'est qu'une fois que nous avons pris en compte toutes les familles du lexique que nous connaissons les patrons

réguliers au niveau de relations sémantiques et le bon niveau d'abstraction. Par exemple, si nous prenons en compte uniquement la famille de POT et celle de BRIQUE et TUILE présentées en (56), qui présentent une structure similaire, nous pourrions être tentés de considérer que la propriété sémantique déterminante pour la structuration de ces familles est le fait qu'elles sont basées sur un lexème qui dénote un objet en terre cuite, mais une fois que nous prenons en compte la famille de BOTTE ou celles de GANT et CARTE en (57), nous voyons qu'elles présentent la même structure et que c'est plutôt la nature d'objet fabriqué de l'élément initial de la famille qui est pertinent. C'est pour cette raison que nous considérons que le bon niveau de généralisation pour l'identification des paradigmes dérivationnels ne peut être établi qu'en prenant en compte toutes les familles du lexique.

- (56) a. {BRIQUE.N, BRIQUETIER.N, BRIQUETERIE.N₁ (usine où l'on fait les briques), BRIQUETERIE.N₂ (activité de fabrication des briques)}
- b. {TUILE.N, TUILIER.N, TUILERIE.N (usine où l'on fait des tuiles)}
- (57) a. {GANT.N, GANTIER.N, GANTERIE.N₁ (activité), GANTERIE.N₂ (lieu)}
- b. {CARTE.N, CARTIER.N, CARTERIE.N₁ (activité), CARTERIE.N₂ (lieu)}

Le LMB que nous proposons en Figure 4.20 est le résultat de l'alignement des relations contenues dans les FMBs décrivant une action/activité de fabrication et de l'abstraction de leur contenu. Les étiquettes ontologiques utilisées dans le LMB en Figure 4.20 représentent le dénominateur commun aux participants qui sont impliqués dans les FMBs que nous avons proposés pour les familles de POT, BOTTE, TOUR et FAÏENCE et qui tient compte également du contenu des familles en (56) et (57). Dans ce LMB :

- l'OBJET FABRIQUÉ est un ARTEFACT ;
- le FABRICANT est une PERSONNE ;
- l'INSTRUMENT utilisé est un ARTEFACT ;
- on utilise des MATÉRIAUX qui correspondent à des MATIÈRES ;
- la fabrication se déroule dans un LIEU DE FABRICATION qui est lui aussi un ARTEFACT ;

- on a des DÉCORATIONS, qui sont aussi des ARTEFACTS qui peuvent être réalisées sur l'objet fabriqué ;

Nous rappelons que dans un *meaning bundle* on peut avoir plusieurs participants qui partagent la même catégorie ontologique (ex. ARTEFACT) et qui sont distingués par rapport au rôle qu'ils jouent dans la situation décrite par le MB (ex. OBJET FABRIQUÉ, INSTRUMENT, LIEU DE FABRICATION, DÉCORATION).

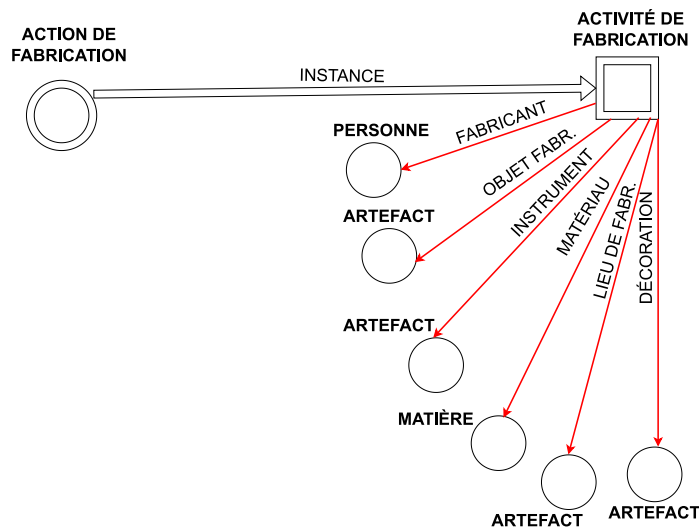


FIGURE 4.20 – LMB décrivant une activité et une action de fabrication.

Pour établir le niveau de généralisation à considérer pour aligner les FMBs de POT et de BOUTEILLE que nous avons proposés précédemment, nous avons besoin également de prendre en compte la totalité des familles contenues dans le lexique. En (58) nous proposons d'autres familles basées sur des récipients qui présentent les mêmes relations que celles que nous avons identifiées dans les FMBs créés à partir des familles de POT et de BOUTEILLE. Nous considérons que le fait d'être basées sur un nom dénotant un récipient est la propriété sémantique pertinente pour la structuration des familles de POT, de BOUTEILLE et de celles en (58).

- (58)
- {CARAFE.N, CARAFER.V}
 - {CAISSE.N, ENCAISSER.V, ENCAISSEUR.N}
 - {TONNEAU.N, ENTONNER.V, ENTONNOIR.N}

Sur cette base, nous proposons le LMB en Figure 4.21, qui décrit un scénario de déplacement d'une entité concrète dans un récipient. Ce LMB implique :

- un AGENT DU DÉPLACEMENT de type PERSONNE ;
- une ENTITÉ DÉPLACÉE de type ENTITÉ CONCRÈTE (ex. plante pour le rempotage, substances pour des mises en bouteille ou en carafe) ;
- une DESTINATION de type RÉCIPIENT ;
- un INSTRUMENT de type ARTEFACT (ex. embouteilleuse, autres machines, entonnoir)

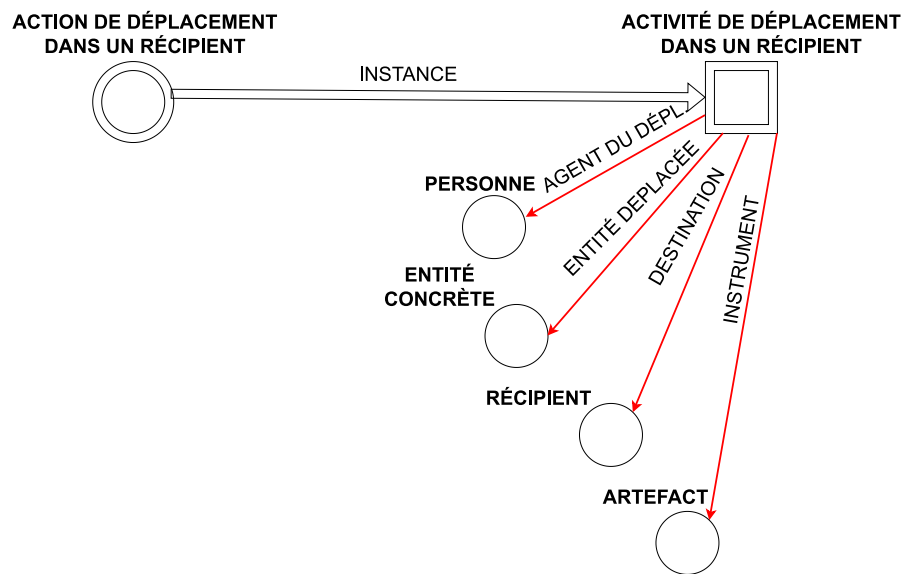


FIGURE 4.21 – LMB décrivant le déplacement d'une entité concrète dans un récipient.

Une fois arrivés à ce stade de la démarche, nous avons obtenu des LMBs qui généralisent le contenu des histoires au niveau de toutes les familles du lexique et qui contiennent des participants interprédictibles qui apparaissent régulièrement ensemble dans les histoires. Dans la Figure 4.22 nous récapitulons la procédure adoptée jusqu'ici pour obtenir les deux LMB que nous venons de présenter et que l'on utilise généralement pour obtenir des LMBs¹⁷.

17. En revanche, nous allons voir qu'en fonction des propriétés sémantiques des lexèmes considérés l'alignement doit être abordé de manière différente (voir Chapitre 5).

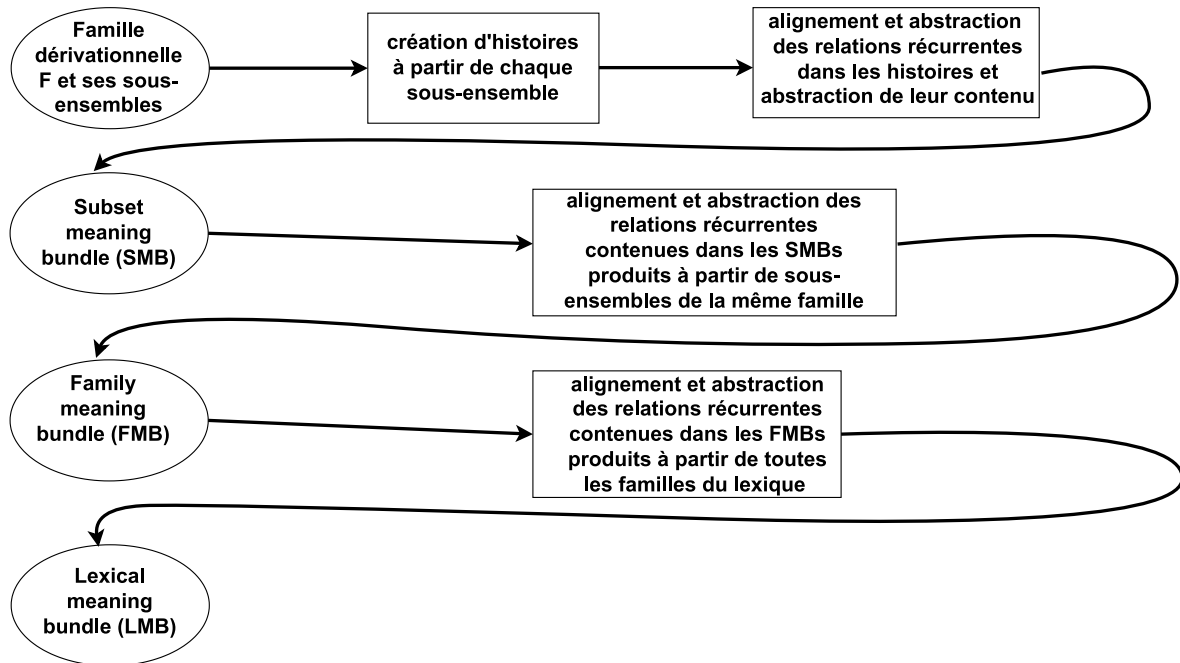


FIGURE 4.22 – Schématisation de la procédure que nous avons adoptée pour obtenir des *lexical meaning bundles*.

4.13 Projection des LMBs sur le lexique

La dernière étape de notre méthode prévoit la projection des LMBs sur les familles dérivationnelles du lexique pour les découper et produire des familles paradigmatiques alignables. Dans cette étape, nous projetons les réseaux à partir desquels nous avons obtenu les LMBs sur les familles présentes dans le lexique construit. Ces familles sont les familles sous forme de liste dont nous sommes partis.

Les deux LMBs que nous avons obtenus se composent d'un volet action et d'un volet activité. Comme nous l'avons vu, le volet action et activité ont la même structure et contiennent les mêmes relations. En outre, une activité est considérée comme une répétition de plusieurs procédés qui sont décrits par des réseaux action. Sur cette base, nous considérons qu'en projetant un réseau activité sur le lexique construit nous projetons déjà tout le contenu sémantique du LMB en termes de participants et de relations.

La projection des LMBs sur les familles complètes est schématisée en Figure 4.23 : elle

permet d'identifier les lexèmes contenus dans les familles complètes qui possèdent les propriétés ontologiques et relationnelles des participants du LMB. Les lignes pointillées symbolisent les correspondances que l'on peut établir grâce à la projection. Cette étape permet de découper la famille de POT dont nous sommes partis : *REMPOTER* et *REMPOTEUR* ne réalisent aucun participant du LMB projeté et donc ne vont pas faire partie de la famille paradigmatique produite à partir d'un LMB de fabrication. D'autre part, on voit que certains participants du LMB de fabrication ne sont pas réalisés dans la famille de POT (ex. *INSTRUMENT*, *LIEU DE FABRICATION*, *MATÉRIAU* et *DÉCORATION*).

Le résultat de la projection est un graphe complet contenant des participants qui sont tous reliés sémantiquement l'un par rapport à l'autre (voir partie inférieure de la Figure 4.23). En outre, la projection permet d'identifier des phénomènes de polysémie : par exemple *POTERIE* réalise à la fois l'activité de fabrication et l'*OBJET FABRIQUÉ*.

Nous effectuons la projection des LMBs sur toutes les familles du lexique pour trouver celles qui contiennent des familles paradigmatiques que l'on peut inscrire dans le LMB projeté. Nous proposons d'aligner les graphes que nous avons obtenus avec la projection des LMBs sur les familles et de les organiser dans un format tabulaire, qui est le format le plus utilisé pour représenter les paradigmes en morphologie (voir Introduction et Chapitre 2). Des exemples de paradigmes obtenus avec la projection du LMB en Figure 4.23 sont proposés dans les Tables 4.1, 4.2 et 4.3.

Les colonnes disponibles pour le paradigme sont déterminées par le scénario que le LMB décrit. Chaque sommet du graphe décrivant le LMB correspond à une colonne dans le paradigme. Chaque colonne dans le paradigme contient une catégorie relationnelle (le rôle sémantique du participant) et une catégorie ontologique, sauf pour l'activité qui est le pivot sémantique du LMB. Dans les paradigmes que nous proposons, une cellule d'un tableau peut être occupée par un seul lexème. Lorsque deux lexèmes jouent le même rôle dans le LMB et appartiennent à la même catégorie syntaxique, nous considérons qu'ils appartiennent à deux familles paradigmatiques distinctes qui peuvent être alignées, de la même manière que ce qui est proposé par Hathout et Namer (2022) dans *ParaDis*. Par exemple, *POTERIE* dénote à la fois l'activité de

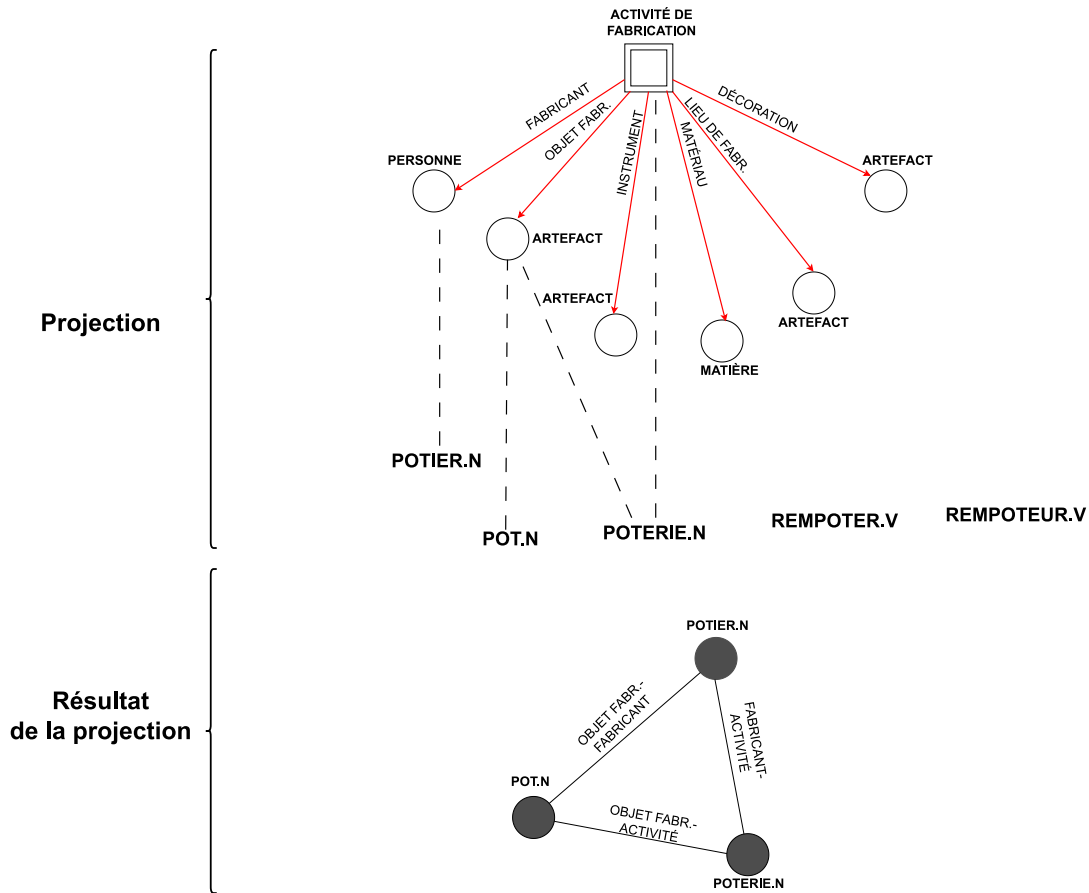


FIGURE 4.23 – Exemple de projection d’un graphe représentant un LMB sur un graphe représentant une famille dérivationnelle. Le résultat de cette projection est un graphe complet qui correspond à une famille paradigmatique.

fabrication ($POTERIE_1$) et un objet fabriqué par un potier ($POTERIE_2$) : $POTERIE_2$ joue le même rôle que POT et ces deux lexèmes s’inscrivent dans deux familles paradigmatiques distinctes qui peuvent être alignées (voir Table 4.1).

Lorsqu’un participant du LMB n’est pas réalisé morphologiquement dans une famille paradigmatique, nous remplissons la cellule correspondante avec ‘-’. On peut avoir également des participants qui ne sont pas du tout réalisés dans les familles paradigmatiques, comme la DÉCORATION : dans ce cas, la colonne correspondante au participant n’est pas présente dans le paradigme.

Le premier paradigme proposé en Table 4.1 contient des familles paradigmatiques qui rentrent

dans un scénario de fabrication. Dans toutes ces familles, l'élément initial est une entité qui jouant le rôle d'OBJET FABRIQUÉ dans un LMB de fabrication. Ce paradigme montre que les familles basées sur des noms dénotant une entité qui joue typiquement le rôle d'OBJET FABRIQUÉ ont une structure assez régulière : elles contiennent typiquement le nom du FABRICANT, le nom de l'ACTIVITÉ qui le produit et le LIEU DE FABRICATION et cela donne un pouvoir prédictif au paradigme. D'autre part, ces familles ne contiennent pas des noms jouant le rôle de MATÉRIAU, INSTRUMENT ou DÉCORATION et donc nous n'incluons pas ces colonnes dans le paradigme.

OBJET FABR. ARTEFACT	FABRICANT PERSONNE	ACTIVITÉ.N	LIEU DE FABR. ARTEFACT
POT	POTIER	POTERIE	POTERIE
POTERIE	POTIER	POTERIE	POTERIE
BOTTE	BOTTIER	BOTTERIE	BOTTERIE
BRIQUE	BRIQUETIER	BRIQUETERIE	BRIQUETERIE
TUILE	TUILIER	-	TUILERIE
GANT	GANTIER	GANTERIE	GANTERIE
CARTE	CARTIER	CARTERIE	CARTERIE

TABLE 4.1 – Paradigmes regroupant un extrait des familles de lexèmes qui interviennent dans une activité de fabrication et qui ont comme élément initial une entité jouant le rôle d'OBJET FABRIQUÉ dans un LMB de fabrication.

Ensuite, nous pouvons identifier un deuxième paradigme qui contient des familles paradigmatiques dont l'élément initial est un nom dénotant une entité qui joue le rôle de MATÉRIAU dans la fabrication. Ces familles sont organisées dans la Table 4.2. Elles sont aussi structurées de manière régulière : elles contiennent le FABRICANT, l'ACTIVITÉ DE FABRICATION qui fait usage du MATÉRIAU et l'OBJET FABRIQUÉ, qui dans les familles en Table 4.2 a toujours la même forme que le nom du MATÉRIAU. Dans la famille de FAÏENCE, le LIEU DE FABRICATION est également réalisé.

En troisième lieu, nous proposons en Table 4.3 un paradigme plus réduit qui contient des familles paradigmatiques dont l'élément initial est un nom dénotant une entité qui intervient en tant qu'INSTRUMENT dans un LMB de fabrication. Les deux familles que nous organisons dans ce paradigme sont la famille de TOUR et la famille de FRAISE que nous avons traitées

MATÉRIAU MATIÈRE	FABRICANT PERSONNE	ACTIVITÉ.N	LIEU DE FABR. ARTEFACT	OBJET FABR. ARTEFACT
FAÏENCE	FAÏENCIER	FAÏENCERIE	FAÏENCERIE	FAÏENCE
FAÏENCE	FAÏENCIER	FAÏENCERIE	FAÏENCERIE	FAÏENCERIE
BRONZE	BRONZIER	BRONZE	-	BRONZE
BRONZE	BRONZIER	BRONZERIE	-	BRONZERIE
CÉRAMIQUE	CÉRAMISTE	CÉRAMIQUE	-	CÉRAMIQUE

TABLE 4.2 – Paradigmes regroupant un extrait des familles de lexèmes qui interviennent dans une activité de fabrication

en Section 4.12. Dans ce paradigme, les familles paradigmatiques réalisent le FABRICANT et l'activité sous forme verbale. Bien entendu, dans ces familles on a également TOURNAGE et FRAISAGE, dont le contenu sémantique est équivalent à celui de TOURNER et FRAISER (nous discutons la relation entre verbes et noms d'action/activité dans le Chapitre 6).

INSTRUMENT ARTEFACT	FABRICANT PERSONNE	ACTIVITÉ.V
TOUR	TOURNEUR	TOURNER
FRAISE	FRAISEUR	FRAISER

TABLE 4.3 – Paradigme regroupant un extrait des familles de lexèmes qui interviennent dans une activité de fabrication et qui ont comme élément initial un nom dénotant un INSTRUMENT utilisé dans la fabrication.

Les trois paradigmes que nous venons de présenter dans les Tables 4.1, 4.2 et 4.3 montrent qu'à partir d'un LMB donné on peut produire plusieurs paradigmes qui se distinguent sur la base de leur élément initial et de la réalisation des rôles : par exemple, le MATÉRIAU est réalisé uniquement dans les familles paradigmatiques en Table 4.2.

Pour sa part, la Table 4.4 présente un extrait des familles paradigmatiques qui interviennent dans une activité de déplacement dans un récipient et qui sont toutes basées sur un nom de récipient. Dans ce cas, nous considérons que la propriété 'récipient' est déterminante pour la structuration des familles paradigmatiques : si la famille est basée sur un nom qui dénote une entité ayant une fonction de RÉCIPIENT, probablement il y aura dans la même famille l'ACTIVITÉ DE DÉPLACEMENT et l'AGENT DU DÉPLACEMENT dans la même famille. En outre, on peut également trouver dans la famille l'INSTRUMENT utilisé dans le déplacement.

DESTINATION RECIPIENT	ACTIVITÉ.V	INSTRUMENT ARTEFACT	AGENT DU DÉP. PERSONNE
POT	REMPOTER	-	REMPOTEUR
BOUTEILLE	EMBOUTEILLER	EMBOUTEILLEUSE	EMBOUTEILLEUR
CAISSE	ENCAISSER	-	ENCAISSEUR
CARAFE	CARAFER	-	-
TONNEAU	ENTONNER	ENTONNOIR	-

TABLE 4.4 – Paradigme dérivationnel regroupant des familles de lexèmes qui interviennent dans une activité de déplacement dans un récipient.

Les deux paradigmes que nous venons de proposer montrent que les paradigmes dérivationnels, comme les *meaning bundles*, peuvent se chevaucher. Cela concerne le lexème POT, que nous retrouvons dans les deux paradigmes que nous venons de proposer avec deux rôles distincts.

4.14 Bilan

Nous avons présenté dans ce chapitre une méthode théorique qui permet d’identifier des paradigmes dérivationnels délimités sur la base de scénarios, que nous avons appelés *meaning bundles*. Nous l’avons illustrée sur un exemple de famille dérivationnelle de départ, à savoir la famille de POT. Nous avons basé notre démarche sur la représentation des relations sémantiques entre les lexèmes au moyen d’histoires. Comme nous l’avons vu, les histoires servent à faire émerger les relations récurrentes que l’on trouve dans le lexique construit et qui structurent les familles.

Pour identifier les relations récurrentes dans les histoires, nous nous sommes basés principalement sur des critères ontologiques et relationnels. L’alignement et l’abstraction du contenu des histoires permettent d’obtenir les MBs. Nous nous sommes servis de trois types de *meaning bundles* que nous avons appelés respectivement SMB, FMB et LMB et qui interviennent dans des étapes distinctes de la procédure. Deux LMBs ont été identifiés (fabrication vs déplacement dans un récipient) et ensuite projetés sur le lexique construit. Cette projection des LMBs sur les graphes représentant les familles dérivationnelles a produit des graphes qui correspondent aux familles paradigmatiques. Nous avons proposé de représenter ces graphes que nous avons

obtenus en utilisant un format tabulaire. Les tables que nous avons proposées contiennent donc des paradigmes dérivationnels. La première ligne des paradigmes dérivationnels que nous proposons contient les informations ontologiques et relationnelles relatives aux participants du LMB. Chaque colonne dans un paradigme dérivationnel contient des lexèmes qui jouent le même rôle dans le scénario.

Nous avons mis en évidence plusieurs propriétés des paradigmes dérivationnels que nous allons retrouver également dans la suite de la thèse : **(a)** les paradigmes dérivationnels sont structurés par des scénarios décrivant des domaines d'activité ; **(b)** les paradigmes dérivationnels peuvent se chevaucher (i.e. partager des lexèmes) ; **(c)** lorsque deux paradigmes se chevauchent, ils dépendent des propriétés sémantiques distinctes d'un même lexème ; **(d)** les propriétés sémantiques des éléments initiaux des familles jouent un rôle central pour la structuration des familles paradigmatiques ; **(e)** les paradigmes dérivationnels présentent des trous lorsqu'un participant du LMB projeté n'est pas réalisé morphologiquement dans une famille paradigmatique ; **(f)** on peut produire plusieurs paradigmes en projetant un LMB sur les familles dérivationnelles du lexique. Ces paradigmes se distinguent sur la base du rôle joué par leur élément initial ; **(g)** on peut obtenir plusieurs familles paradigmatiques alignables dans le même paradigme à partir de la même famille (voir Tables 4.1 et 4.2).

Dans la suite de cette thèse, nous étendons l'inventaire des LMBs que nous pouvons produire à partir des familles dérivationnelles. Nous allons considérer des familles qui contiennent des noms appartenant à d'autres catégories (i.e. animaux et fruits) pour déterminer dans quels LMBs et quels paradigmes on peut les inscrire (c.f. Chapitre 5). En outre, au Chapitre 6 nous considérons un ensemble de dérivés comme les adjectifs relationnels, la relation entre verbes et noms d'action, la relation masculin-féminin chez les noms d'humain et les préfixations en *re-*. Nous considérons que ces dérivés s'inscrivent dans des relations qui ne peuvent pas être révélées par les histoires et nécessitent d'un traitement spécifique.

Chapitre 5

Propriétés sémantiques des lexèmes et des paradigmes

Sommaire

5.1	Introduction	146
5.2	Relations sociales et comportement animal	147
5.2.1	Comportement social des animaux	149
5.2.2	Abri et groupe social	155
5.2.3	Proies et prédateurs	159
5.3	Animaux et activités humaines	163
5.3.1	Pêche	163
5.3.2	Chasse et animaux “nuisibles”	169
5.3.3	Élevage	176
5.3.4	Animaux en tant qu’objet d’étude	182
5.4	Bilan intermédiaire	185
5.5	Fruits	187
5.5.1	Production fruiticole	187
5.5.2	Transport spécialisé	191
5.5.3	Cuisine et recettes	195
5.6	Exemples de relation entre LMBs	199

5.7 Bilan	202
----------------------------	------------

5.1 Introduction

Dans le Chapitre 4 nous avons proposé une démarche théorique pour résoudre le problème du découpage des familles dérivationnelles en familles paradigmatiques alignables. Notre démarche est guidée par les propriétés sémantiques des lexèmes contenus dans les familles dérivationnelles et par les histoires dans lesquelles on peut inscrire ces lexèmes. Nous avons montré que les *meaning bundles* et les paradigmes dérivationnels peuvent se chevaucher et sélectionner des propriétés sémantiques diverses du même lexème.

Dans ce chapitre, nous approfondissons la question du rôle des propriétés sémantiques des lexèmes dans la structuration des paradigmes en dérivation. Nous considérons des familles contenant des lexèmes appartenant à deux catégories ontologiques : les noms d’animaux (N_{anim}) et les noms de fruits (N_{fru}). Les relations dérivationnelles impliquant des N_{anim} ont fait l’objet de travaux récents que nous utilisons comme point de départ pour notre discussion (Boyé et Schalchli, 2017; Fradin, 2020b). Il en va de même pour les familles basées sur N_{fru} , qui ont été également étudiées dans des travaux récents dans une perspective paradigmatique. Par exemple, selon Fradin (2020a), les familles basées sur des noms de fruit donnent lieu à des paradigmes qui sont uniquement basés sur leurs propriétés ontologiques et qui ne dépendent pas d’une action ou d’une activité (voir Section 2.4.1), contrairement aux paradigmes que nous avons vus au Chapitre 4, qui pour leur part s’inscrivent dans des activités humaines. Cela suggère que les *meaning bundles* que nous pouvons obtenir à partir des familles basées sur N_{fru} contiennent des relations de nature différente par rapport aux LMBs identifiés au Chapitre 4. Dans ce chapitre, nous vérifions cette hypothèse (voir Section 5.5).

Plus généralement, l’objectif de ce chapitre est d’approfondir la question du rôle des propriétés sémantiques des lexèmes pour la structuration des paradigmes. Nous voulons vérifier dans quelle mesure les catégories ontologiques ANIMAL et FRUIT sont pertinentes pour la structuration des paradigmes, déterminer les *meaning bundles* dans lesquels ils s’inscrivent et

quels sont les paradigmes dérivationnels que l'on obtient en projetant ces *meaning bundles* sur le lexique construit.

Les LMBs que nous allons proposer dans ce chapitre permettent de découper en familles paradigmatiques des familles basées sur N_{anim} comme celles en (59). Comme le montre la famille en (59c), dans ce chapitre nous allons également prendre en compte la possibilité d'intégrer des composés savants tels que les *Xculteur* et *Xculture* aux paradigmes dérivationnels.

- (59) a. {LOUP.N, LOUVE.N, LOUVETEAU.N, LOUVETIER.N (chasseur de loups), LOUVETIERIE.N (activité de chasse des loups)}
- b. {FOURMI.N, FOURMILIER.N, FOURMILIÈRE.N}
- c. {CREVETTE.N, CREVETTIER.N, CREVETTICULTEUR.N (éleveur de crevettes), CREVETTICULTURE.N (élevage des crevettes)}

En outre, nous allons montrer comment on peut produire des familles paradigmatiques qui rentrent dans des paradigmes distincts sur la base des différents sens qui sont associés à BANANIER dans la famille en (60) (i.e. 'arbre à bananes', 'gâteau à base de bananes', 'cargo qui transporte des bananes').

- (60) {BANANE.N, BANANIER.N, BANANERAIE.N}

5.2 Relations sociales et comportement animal

Nous nous intéressons en premier lieu aux familles dérivationnelles basées sur N_{anim} . Les référents de ces noms rentrent dans la catégorie des êtres vivants doués d'intentionnalité (Flaux et Van de Velde, 2000; Haas *et al.*, 2023). Parmi les animaux, Fradin (2020b) identifie trois sous-catégories :

- animaux avec lesquels l'être humain a des interactions quotidiennes, comme les animaux domestiques (ex. vache);
- animaux sauvages avec lesquels l'homme est en relation dans le cadre de la chasse ou de la pêche (ex. loup);

- animaux sauvages dont le contact avec l’homme est réduit ou non existant (ex. poisson-lanterne).

Dans ce chapitre, nous prenons en compte des sous-ensembles de familles contenant des N_{anim} appartenant aux deux premières catégories d’animaux identifiées par Fradin : les animaux sauvages qui sont en relation avec l’homme et les animaux domestiques.

Concernant les lexèmes dérivés de N_{anim} que l’on trouve dans le lexique construit, Fradin (2020b) propose une distinction entre deux types de relations morphosémantiques :

- les relations morphosémantiques internes à l’espèce animale, qui comprennent les relations entre l’animal adulte et son petit (ex. OURS - OURSON), entre l’animal et son abri (ex. TERMITE - TERMITIÈRE) et entre l’animal et une de ses caractéristiques physiques (ex. TIGRE - TIGRÉ) ;
- les relations morphosémantiques externes à l’espèce animale, qui comprennent les relations entre le nom dénotant l’espèce animale et un nom d’humain (ex. LOUP - LOUVETIER) ou entre l’espèce animale et une autre espèce animale (ex. FOURMI - FOURMILIER).

Nous prenons en compte des dérivés qui appartiennent à ces deux classes : nous procédons en considérant d’abord les dérivés qui n’impliquent pas d’activité humaine et qui concernent la vie des animaux, leurs relations sociales et leur comportement (Section 5.2) et dans un deuxième temps en Section 5.3 nous prenons en compte des familles contenant des lexèmes qui renvoient à des activités humaines variées où l’animal est un “protagoniste”. Nous faisons ce choix pour isoler les éléments des familles de N_{anim} qui ne s’inscrivent pas dans le cadre d’une activité humaine et voir si les paradigmes que l’on arrive à identifier à partir de ces lexèmes sont uniquement déterminés par les propriétés ontologiques associés aux N_{anim} , de manière similaire à ce que Fradin (2021) suggère pour les familles basées sur N_{fru} . Pour sa part, le traitement de dérivés qui s’inscrivent dans des activités humaines où l’animal est un protagoniste nous permettra d’obtenir des nouveaux réseaux de type activité relevant de domaines d’activité différents. Cela est en ligne avec la proposition de Fradin (2021, p.23) « [...] il serait intéressant de voir si des réseaux (plus) stables ne peuvent pas être établis sur la base de domaines ou sous-domaines

d'activité, ce qui revient à prendre en compte l'ancrage social de ces réseaux. »¹⁸

5.2.1 Comportement social des animaux

Nous prenons comme premier exemple la famille partielle de LOUP proposée en (61). Comme nous venons de l'annoncer, nous excluons les dérivés qui dénotent des humains ou des activités humaines. Nous allons donc prendre en compte LOUVETIER (chasseur de loups) et LOUVETERIE (chasse aux loups) en Section 5.3.

(61) {LOUP.N, LOUVE.N, LOUVETEAU.N}¹⁹.

Nous considérons que les relations entre les lexèmes en (61) sont internes à l'espèce animale des loups : le nom LOUP dénote à la fois l'espèce animale (i.e. *canis lupus*) et un individu mâle appartenant à la même espèce que celle des référents des noms LOUVE et LOUVETEAU. Les histoires produites à partir de la famille en (61) présentent une difficulté : les référents de LOUVE et LOUVETEAU interviennent dans des situations très variées, car une louve peut donner la vie à un louveteau, le protéger, l'allaiter, lui apprendre à chasser, etc. Par conséquent, lorsqu'on produit des histoires à partir d'un couple comme {LOUVE, LOUVETEAU} on obtient de manière régulière des graphes contenant des événements et des relations de nature différente. Pour cette raison, pour abstraire le contenu de ces histoires et déterminer la relation existant entre LOUVE et LOUVETEAU, nous ne pouvons pas faire appel à l'alignement que nous avons décrit pour les SMBs obtenus à partir de couples comme {POT, POTIER} au Chapitre 4, où la grande majorité des histoires décrit des actions ou des activités de même nature.

Nous considérons que tous les événements qui ne sont pas épisodiques et apparaissent régulièrement dans les histoires construites à partir du couple {LOUVE, LOUVETEAU} contribuent à caractériser la relation sémantique reliant ces deux lexèmes. Cela veut dire que cette relation est un type d'objet différent des relations argumentales contenues dans les *meaning bundles* de

18. « [...] *it would be worth to see whether (more) stable networks cannot be established on the basis of activity domains or sub-domains, which amounts to take into account the social anchoring of these networks.* » dans la version originale

19. Pour simplifier l'illustration, dans la famille partielle de départ en (61) nous n'avons pas pris en compte non plus le nom LOUVARD, qui dénote un jeune loup.

type action et activité que nous avons introduits au Chapitre 4. En outre, cette relation ne peut pas non plus être décrite au moyen de relations lexicales classiques comme l'hyponymie. Nous schématisons la relation entre LOUVE et LOUVETEAU en Figure 5.1 : cette relation est représentée par une arête jaune qui représente l'abstraction d'un ensemble d'événements de nature diverse (ex. protéger, nourrir, mettre bas). Pour leur part, ces événements sont représentés par des réseaux action comportant des rôles sémantiques divers : par exemple, par rapport à l'action de nourrir effectuée par la louve, le louveteau est un BÉNÉFICIAIRE, alors que par rapport à l'action de mettre bas effectuée par la louve, le louveteau est considéré comme étant un PATIENT (ces deux rôles ont été présentés au Chapitre 3). Ce qui permet d'assurer la généralisation du contenu des réseaux action en Figure 5.1 est le fait que tous ces graphes décrivent des actions régulièrement présentes dans les histoires qui impliquent une louve et un louveteau.

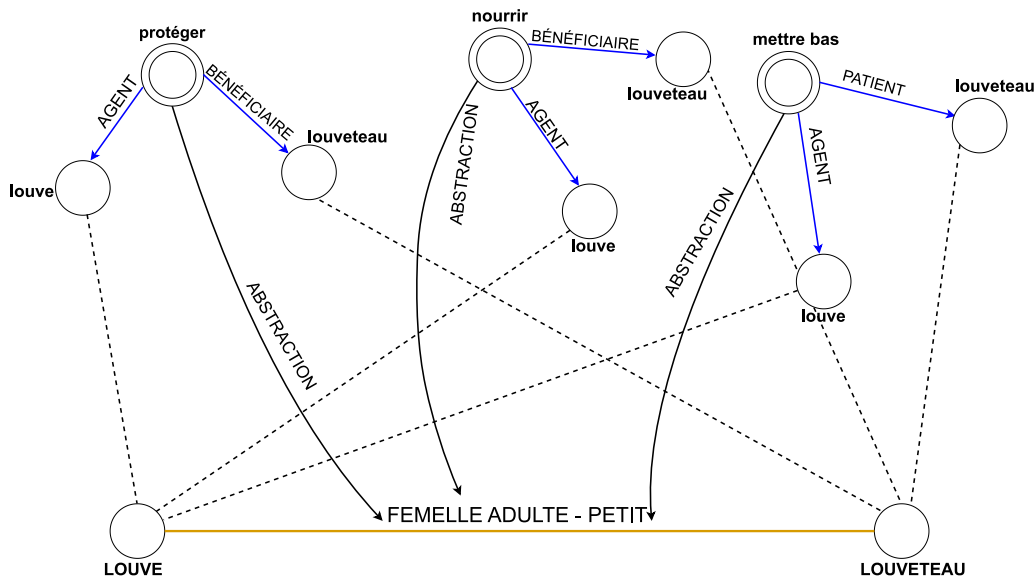


FIGURE 5.1 – Représentation de la relation entre LOUVE et LOUVETEAU. Cette relation, représentée par une arête jaune, est une abstraction d'un ensemble d'événements de nature variée dans lesquels une louve et un louveteau interviennent. Cette relation marche dans les deux sens et n'est pas orientée. Pour la caractériser, nous utilisons l'étiquette FEMELLE ADULTE-PETIT qui abstrait des événements de nature diverse. Les lignes pointillées montrent que ce qui est régulier dans les réseaux action est la présence d'une louve et d'un louveteau.

Nous faisons la même proposition pour caractériser la relation qui existe entre LOUP et

LOUVE et celle qui relie LOUP à LOUVETEAU. Pour caractériser ces deux relations, les locuteurs peuvent également évoquer de manière régulière diverses situations : un loup peut s'accoupler avec une louve, ils peuvent chasser ensemble ou ils peuvent guider la meute s'il s'agit d'un mâle et d'une femelle alpha. En outre, un loup mâle adulte peut protéger un louveteau ou lui apprendre comment chasser. Il en va de même lorsqu'on considère le triplet {LOUP, LOUVE, LOUVETEAU} et on essaie de construire des histoires qui impliquent en même temps un loup, une louve et un louveteau : dans ce cas aussi, on aura régulièrement des histoires de nature variée.

Sur cette base, nous considérons que les relations reliant les lexèmes en (61) dont nous sommes partis structurent un FMB décrivant les relations sociales entre des exemplaires de loups : naissance, croissance, âge adulte, etc. Ce FMB est obtenu en alignant et en regroupant les relations contenues dans les SMBs que l'on peut obtenir à partir des sous-ensembles en (62). Le FMB décrivant les relations sociales entre loups est proposé en Figure 5.2 : contrairement aux FMBs du Chapitre 4 il n'est pas structuré par des relations de nature argumentale. En outre, il ne présente pas de pivot sémantique (contrairement au réseau activité), car il se présente sous forme de graphe acentrique.

- (62) a. {LOUVE, LOUVETEAU}
 b. {LOUP, LOUVE}
 c. {LOUP, LOUVETEAU}
 d. {LOUP, LOUVE, LOUVETEAU}

Une fois obtenu un FMB à partir de la famille partielle de LOUP en (61), notre démarche propose d'aligner ce FMB avec d'autres FMBs obtenus à partir d'autres familles dérivationnelles qui contiennent des relations MÂLE-FEMELLE-PETIT alignables. Établir quelles sont les espèces animales pertinentes n'est pas trivial : les animaux ont des comportements et des organisations sociales extrêmement variés et complexes et le rôle du mâle, de la femelle et du petit varient beaucoup en fonction de l'espèce animale considérée. Un premier choix pertinent pour trouver des familles qui peuvent produire des FMBs alignables avec celui en Figure 5.2 est de

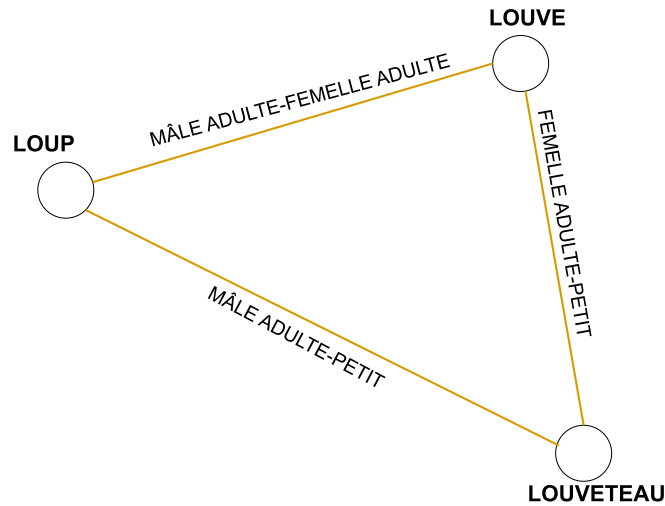


FIGURE 5.2 – FMB décrivant les relations sociales entre exemplaires d’un loup. Contrairement aux FMB décrits dans le Chapitre 4, ce FMB est structuré par un graphe acentrique contenant des relations non orientées qui ne sont pas de nature argumentale ni font partie des relations lexicales classiques comme l’hyponymie. Ces relations sont également présentes dans l’inventaire de relations lexicales proposé par Nutter (1989, p.13)

prendre en compte uniquement des animaux sauvages. Nous excluons les animaux domestiques pour deux raisons : une raison liée à la réalité extralinguistique et une raison morphologique. Sur le plan extralinguistique, nous considérons que chez les animaux domestiques les interactions entre mâle, femelle et petit ne sont pas les mêmes que celles qu’on a chez les animaux sauvages : par exemple, les vaches sont souvent fécondées artificiellement et n’ont pas d’interaction avec les taureaux. Nous considérons que les noms d’animaux domestiques ont des propriétés “ontologiques” liées à leur utilisation dans notre société qui les rapprochent à des produits : on les utilise pour la viande, le lait, la laine, etc. En outre, d’un point de vue morphologique, chez les N_{anim} dénotant des animaux domestiques on trouve plusieurs cas où le nom dénotant le mâle d’une espèce animale n’appartient pas à la même famille que les noms dénotant la femelle ou le petit, comme dans le cas des ensembles {MOUTON, BÉLIER, BREBIS, AGNEAU} et {TAUREAU, BŒUF, VACHE, VEAU, GÉNISSE} (Boyé et Schalchli, 2017).

D’autre part, nous considérons qu’une famille comme celle de LION, dont un sous-ensemble est proposé en (63), est pertinente pour produire un FMB qui peut être aligné avec le FMB

proposé en Figure 5.2. Cette famille contient LION (exemplaire mâle adulte de l'espèce *panthera leo*), LIONNE (femelle adulte de la même espèce) et LIONCEAU (petit de la même espèce).

(63) {LION.N, LIONNE.N, LIONCEAU.N}

Nous considérons que les relations reliant les lexèmes du triplet {LION, LIONNE, LIONCEAU} peuvent être caractérisées par des histoires de nature similaire à celles qui caractérisent les relations reliant {LOUP, LOUVE, LOUVETEAU} : pour les deux espèces, la femelle adulte donne la vie au petit, le protège, le nourrit et l'éduque. En outre, pour les deux espèces, la relation entre mâle et femelle adulte peut être caractérisée par des histoires qui décrivent leur accouplement ou leur rôle dans le groupe social. Pour la relation entre mâle adulte et petit, nous considérons que pour les deux espèces les mâles adultes peuvent apprendre aux petits comment trouver de la nourriture, les protéger ou leur apprendre comment diriger le groupe social. Sur cette base, nous considérons qu'à partir des familles de LOUP et LION on obtient deux FMBs qui peuvent être alignés selon des critères ontologiques et relationnels. Nous précisons à nouveau qu'il s'agit d'une approximation visant à réaliser un alignement relationnel assez stable entre réseaux sémantiques décrivant le comportement social d'espèces animales distinctes.

En outre, nous considérons que les deux FMBs que l'on peut produire à partir de la famille de LOUP et LION peuvent être alignés à d'autres FMBs produits à partir d'autres familles dérivationnelles construites sur N_{anim} dénotant des animaux sauvages dont l'organisation sociale a des traits en commun avec celles des loups et des lions, comme ÉLÉPHANT, RENARD et ZÈBRE (64). Ces animaux sont tous des animaux sociaux qui vivent en groupes hiérarchisés qui sont généralement "commandés" par un mâle dominant. Le cas de ZÈBRE montre que les rôles MÂLE, FEMELLE et PETIT ne sont pas systématiquement distingués sur le plan morphologique dans la famille de départ, car ZÈBRE peut dénoter un exemplaire mâle adulte ou une femelle adulte.

(64) a. {ÉLÉPHANT.N, ÉLÉPHANTE.N, ÉLÉPHANTEAU.N}
 b. {RENARD.N, RENARDE.N, RENARDEAU.N}
 c. {ZÈBRE.N₁ (mâle adulte), ZÈBRE.N₂ (femelle adulte), ZÉBREAU.N }

En Figure 5.3 nous représentons l’alignement entre les FMBs produits à partir des familles de LOUP et LION et l’abstraction de leur contenu, qui prend en compte également de tous les autres FMBs produits à partir de familles qui contiennent les sous-ensembles en (64). Au cours de l’abstraction, nous remplaçons les participants des FMBs par des étiquettes qui expriment leur dénominateur commun, à savoir ANIMAL SAUVAGE MÂLE, ANIMAL SAUVAGE FEMELLE et ANIMAL SAUVAGE PETIT. L’alignement et l’abstraction schématisés en Figure 5.3 produisent un LMB. Ce LMB est pertinent uniquement pour certaines espèces d’animaux sauvages qui s’organisent en groupes sociaux hiérarchisés : nous considérons que la catégorie ontologique ANIMAL est peu efficace pour déterminer des alignements de relations sémantiques stables entre espèces animales différentes.

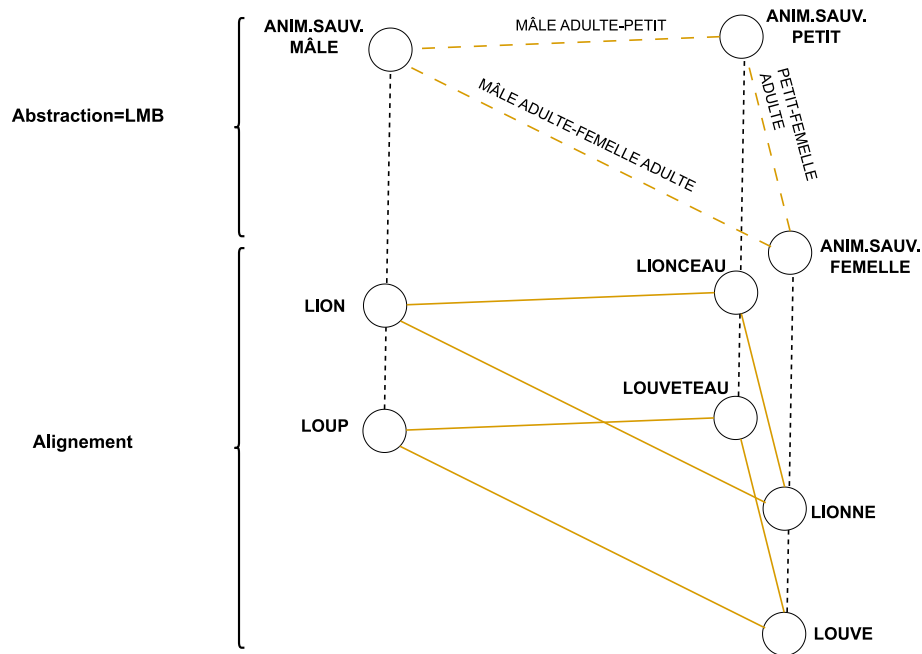


FIGURE 5.3 – Alignement et abstraction de FMBs décrivant les relations entre exemplaires d’une même espèce animale sauvage. Cette abstraction produit un LMB qui est pertinent uniquement pour les animaux sauvages qui s’organisent en groupes sociaux.

La projection du LMB que nous proposons en Figure 5.3 sur le lexique construit nous permet d’obtenir des familles paradigmatiques que nous organisons dans le paradigme dérivationnel en Table 5.1. Contrairement aux paradigmes que nous avons proposés au Chapitre 4, ce paradigme

ne contient pas une double étiquette relationnelle et ontologique dans la première ligne. Cela s’explique parce que dans le LMB en Figure 5.3 nous n’avons pas utilisé des “vrais” rôles comme FABRICANT ou INSTRUMENT (voir Chapitre 4) qui sont attribués par une action ou une activité : nous nous sommes limités à étiqueter les relations par rapport aux deux sommets qu’elles réunissent et à leurs étiquettes ontologiques (ex. MÂLE ADULTE-FEMELLE ADULTE).

ANIM.SAUV.MÂLE	ANIM.SAUV.FEMELLE	ANIM.SAUV.PETIT
LOUP	LOUVE	LOUVETEAU
LION	LIONNE	LIONCEAU
ÉLÉPHANT	ÉLÉPHANTE	ÉLÉPHANTEAU
RENARD	RENARDE	RENARDEAU
ZÈBRE	ZÈBRE	ZÉBREAU

TABLE 5.1 – Paradigme dérivationnel regroupant des familles contenant le mâle, la femelle et le petit d’un animal sauvage social. Chaque colonne correspond à un sommet du LMB en Figure 5.3 et chaque ligne correspond à une famille paradigmatique.

5.2.2 Abri et groupe social

Dans certaines familles contenant de N_{anim} sauvages, on trouve également des noms dénotant des lieux. Ces “abris” peuvent être construits ou non par l’animal et correspondre à son habitat naturel (ex. FOURMILIÈRE, GUÊPIER) ou être des sites de nidification (ex. HÉRONNIÈRE). Nous prenons comme exemple de départ le couple en (65) en laissant de côté FOURMILIER (i.e. animal prédateur des fourmis), qui fait partie de la même famille dérivationnelle et que nous reprendrons dans la Section 5.2.3.

$$(65) \quad \{\text{FOURMI.N, FOURMILIÈRE.N}\}$$

De manière similaire à ce que nous avons vu pour les relations MÂLE-FEMELLE-PETIT dans la section précédente, nous considérons que la relation entre FOURMI et FOURMILIÈRE peut être décrite par des histoires décrivant de diverses situations qui impliquent une fourmi et une fourmilière. Par exemple, on peut avoir des histoires récurrentes dans lesquelles des fourmis construisent une fourmilière, d’autres dans lesquelles les fourmis peuplent une fourmilière ou

d'autres encore qui décrivent leur vie dans la fourmilière. En outre, la fourmi est un animal qui vit en colonie et une fourmilière n'est pas *a priori* peuplée par une seule fourmi, mais plutôt par une colonie. Par ailleurs, FOURMILIÈRE est également un nom collectif dénotant un groupe de fourmis qui vit dans le même nid : il est compatible avec des prédicats de groupe comme *se disperser* (66) (Dubois et Dubois-Charlier, 1996; Huyghe, 2015).

(66) *Toute la fourmilière s'est dispersée.*

Comme les fourmis vivent principalement en colonie, nous considérons que lorsqu'on a une fourmilière (nid) on a une colonie de fourmis qui y vivent. Cela signifie que la colonie de fourmis est interprétable par rapport à la fourmilière. Sur cette base, nous considérons qu'à partir de la famille de FOURMI, qui contient le couple {FOURMI, FOURMILIÈRE} en (65) nous pouvons produire un FMB à trois sommets contenant la fourmi en tant qu'individu, la colonie dont la fourmi fait partie et la fourmilière que la colonie de fourmis peuple. Ce FMB est schématisé en Figure 5.4. Comme pour les FMBs décrivant les interactions sociales entre les exemplaires d'une espèce animale, ce FMB contient des relations qui marchent dans les deux sens et qui ne sont pas de nature argumentale.

Pour typer les relations entre les concepts dans le FMB en Figure 5.4 nous utilisons les étiquettes relationnelles ANIMAL-NID (qui est présente deux fois) et INDIVIDU-GROUPE D'ANIMAUX. De la même manière que pour les FMBs que nous avons proposés en Section 5.2, ces étiquettes ne correspondent pas à des vrais rôles, car elles ne sont pas attribuées par une action ou une activité, mais se limitent à caractériser la relation sur la base des deux concepts que ces relations connectent dans le FMB.

Il se pose la question de trouver d'autres FMBs qui peuvent être alignés avec celui en Figure 5.4 où l'animal en tant qu'individu, son abri et son groupe social sont interprétables. Nous considérons que le nombre de familles qui peuvent produire des FMBs "compatibles" avec celui en Figure 5.4 est très réduit. Par exemple, dans la famille de HÉRON on a HÉRONNIÈRE, qui dénote un lieu où les hérons se retirent et font leurs petits, mais les hérons utilisent cet abri uniquement pour se reproduire et pour nidifier, contrairement aux fourmis, qui vivent la plupart

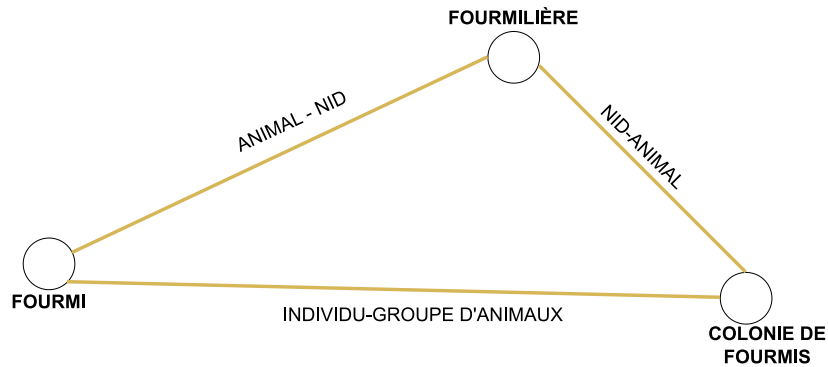


FIGURE 5.4 – FMB produit à partir de la famille de **FOURMI**, qui à son tour contient le couple $\{\text{FOURMI}, \text{FOURMILIÈRE}\}$. Il décrit la relation entre une fourmi en tant qu'individu, sa colonie et la fourmilière où ces fourmis vivent.

de leur vie dans les fourmilières. En outre, un héron peut avoir une héronnière sans faire partie d'une colonie ou d'un groupe social nombreux. La même considération vaut pour la famille de **BLAIREAU**, qui contient le couple $\{\text{BLAIREAU.N}, \text{BLAIREAUTIÈRE.N}\}$ (réseau de galeries où vit le blaireau) : un blaireau peut avoir une tanière sans faire partie d'un groupe social. Cela signifie que le groupe social et l'abri ne sont pas interprédictibles.

D'autre part, nous considérons que les familles qui peuvent produire des FMBs qui contiennent des relations alignables à celles en Figure 5.4 sont basées sur un N_{anim} dénotant des insectes qui vivent en colonie, comme la famille de **TERMITE** et de **GUÊPE**. Pour ces animaux, le **NID** et la **COLONIE** sont interprédictibles, car une **FOURMILIÈRE**, une **TERMITIÈRE** et un **GUÊPIER** sont généralement habités par des colonies et non par des individus. Sur cette base, si on a un nid d'insectes dans une histoire, il est probable que l'histoire parle aussi de la colonie qui l'habite. En outre, de la même manière que **FOURMILIÈRE**, **TERMITIÈRE** et **GUÊPIER** peuvent également dénoter la colonie d'insectes qui habite le nid (68).

- (67) a. $\{\text{TERMITE.N}, \text{TERMITIÈRE.N}_1 \text{ (nid)}, \text{TERMITIÈRE.N}_2 \text{ (colonie)}\}$
 b. $\{\text{GUÊPE.N}, \text{GUÊPIER.N}_1 \text{ (nid)}, \text{GUÊPIER.N}_2 \text{ (colonie)}\}$

- (68) a. *Toute la termitière s'est dispersée.*
 b. *Tout le guêpier s'est dispersé.*

Sur la base du contenu des sous-ensembles en (67) et du sous-ensemble $\{\text{FOURMI}, \text{FOURMILIÈRE}_1, \text{FOURMILIÈRE}_2\}$ dont nous sommes partis, nous considérons que nous pouvons produire un LMB qui décrit la relation entre un insecte en tant qu'individu, sa colonie et le nid dans lequel la colonie vit. Ce LMB est proposé en Figure 5.5. Le LMB en Figure 5.5 décrit donc des régularités sémantiques très spécifiques qui sont propres aux comportements des insectes qui vivent en colonies et qui ne sont pas valables pour toutes les espèces animales.

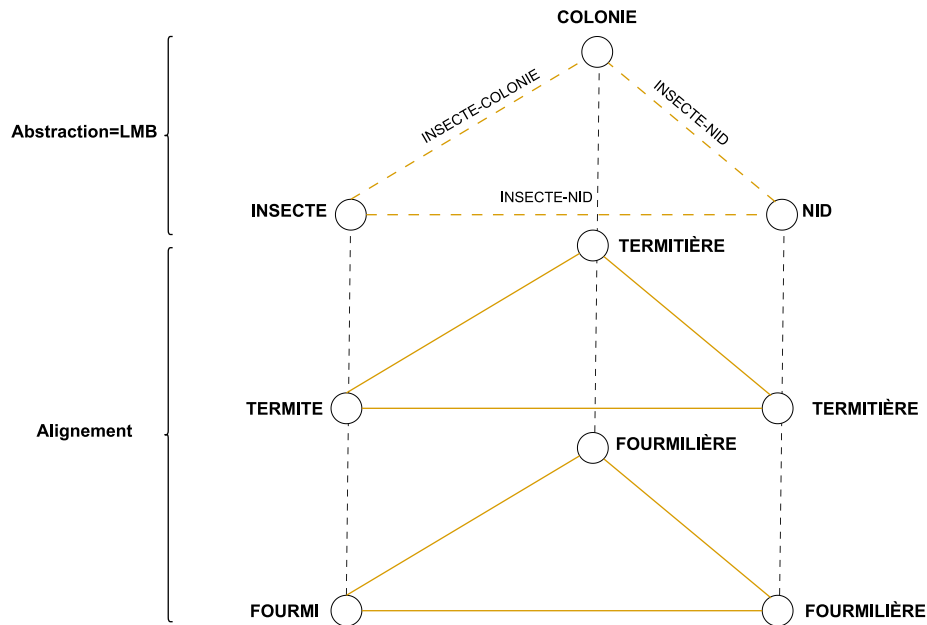


FIGURE 5.5 – LMB structuré par les relations existant entre un insecte, sa colonie et son nid. Ce LMB est pertinent uniquement pour les insectes vivant en colonies et décrit des régularités sémantiques très spécifiques. Les étiquettes ontologiques que nous attribuons aux sommets sont également très spécifiques : nous utilisons les étiquettes INSECTE, COLONIE et NID. Les relations sont déterminées sur la base des concepts qu'elles relient : on a INSECTE-COLONIE et INSECTE-NID.

La projection du LMB proposé en Figure 5.5 sur les familles dérivationnelles du lexique construit donne le paradigme à trois colonnes en Table 5.2. Comme on peut le voir, ce paradigme permet de réunir un nombre très réduit de familles sur le plan sémantique, à savoir les familles basées sur des N_{anim} dénotant des insectes qui vivent en colonies. Cela montre une autre propriété des paradigmes dérivationnels basés sur les relations sémantiques : ils sont structurés parfois par des régularités sémantiques très spécifiques (qui assurent un alignement stable entre les FMBs)

et ils réunissent très peu de familles, comme nous avons pu le voir dans le paradigme précédent en Table 5.1 et encore plus dans le paradigme en Table 5.2.

ANIMAL	NID	COLONIE
FOURMI	FOURMILIÈRE	FOURMILIÈRE
TERMITE	TERMITIÈRE	TERMITIÈRE
GUÊPE	GUÊPIER	GUÊPIER

TABLE 5.2 – Paradigme structuré par les relations sémantiques existant entre un insecte, son abri et sa colonie.

5.2.3 Proies et prédateurs

Nous reprenons la famille de FOURMI pour considérer la relation entre FOURMI et FOURMILIER (69). Contrairement aux relations que nous avons prises en compte dans les deux sections précédentes, la relation entre FOURMI et FOURMILIER est à considérer comme externe à l'espèce animale. Pour revenir à la distinction faite par Fradin (2020b), elle s'établit entre deux espèces animales distinctes : le fourmilier (espèce *Myrmecophaga tridactyla*) est le prédateur de la fourmi (espèces de l'ordre des *formicidae*).

(69) {FOURMI.N, FOURMILIER.N}

Contrairement aux relations que nous avons caractérisées en Section 5.2.1 et 5.2.2, nous considérons que les histoires produites à partir de la famille de FOURMI, dont ce couple fait partie, décrivent régulièrement une action, à savoir une action de chasse (ou plutôt prédation) de la part du fourmilier où la fourmi joue le rôle de la proie. Un exemple d'histoire est fourni en (70).

(70) *Le fourmilier, flairant l'air, était en pleine chasse, traquant la fourmi rapidement. Malgré sa vitesse, la fourmi ne put échapper à sa langue agile.*

L'action de prédation de la part du fourmilier dans les histoires peut aussi ne pas être délimitée temporellement : nous pouvons considérer que les histoires peuvent également décrire

une sorte d'activité de prédation (dans le sens réitératif et régulier du terme) dans laquelle le fourmilier est le PRÉDATEUR habituel et la fourmi est la PROIE habituelle. Sur cette base, nous considérons qu'à partir de la famille de FOURMI, qui comprend le couple {FOURMI.N, FOURMILIER.N}, nous pouvons produire un FMB comme celui en Figure 5.6. Ce FMB décrit une activité de chasse réalisée par le fourmilier, qui joue le rôle de PRÉDATEUR et où la fourmi joue le rôle de la PROIE. Ce FMB n'est pas connexe à celui qui décrit l'organisation en colonie des fourmis, que nous avons pu voir en Section 5.2.2, car sur le plan sémantique un fourmilier n'est pas interprétable par rapport à une fourmière.

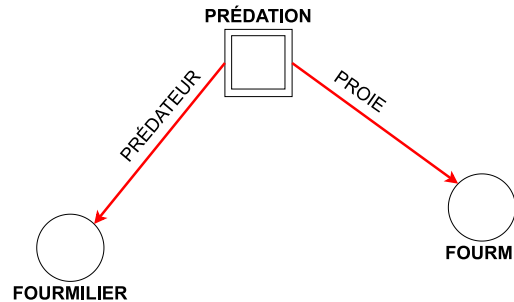


FIGURE 5.6 – FMB décrivant la relation de type PROIE-PRÉDATEUR entre un fourmilier et une fourmi.

Les propriétés sémantiques de FOURMI et FOURMILIER qui sont pertinentes pour le FMB en Figure 5.6 sont associées à leurs rôles respectifs de PROIE et de PRÉDATEUR dans le scénario de prédation qui les implique. Sur cette base, pour trouver d'autres FMBs qui peuvent être alignés avec celui en Figure 5.6 nous devons chercher des FMBs qui décrivent des scénarios de prédation impliquant un animal sauvage prédateur et une proie. À partir de quelles autres familles peut-on produire des FMBs similaires ? Tout en n'étant pas nombreuses, on trouve dans le lexique construit des familles contenant des N_{anim} reliés morphologiquement en synchronie au nom de leur proie comme HUÎTRIER, SERPENTAIRE, CRABIER et GUÊPIER. Les définitions données par le dictionnaire *TLFi* de ces N_{anim} sont fournies en (71).

- (71) a. SERPENTAIRE : « *Rapace diurne au plumage gris, noir et blanc dont le mâle porte sur la tête une huppe érectile, se nourrissant essentiellement de reptiles.* »
- b. CRABIER : « *Héron mangeur de crabes* ».

- c. HUITRIER : « Oiseau échassier au plumage noir et blanc, au bec rouge-orange qui vit le long des côtes et se nourrit essentiellement d’huîtres. »
- d. GUÊPIER : « Nom normalisé donné à trois genres d’oiseaux percheurs de la famille des méropidés [...] dont le régime insectivore se compose d’odonates et d’hyménoptères ».

Nous considérons qu’à partir de couples comme {SERPENT.N, SERPENTAIRE.N}, {CRABE.N, CRABIER.N}, {HUITRE.N, HUITRIER.N} et {GUÊPE.N, GUÊPIER.N (OISEAU)} on peut produire des histoires décrivant des scénarios de prédation et que donc on peut produire des FMBs similaires à celui qui a été proposé en Figure 5.6. Sur cette base, nous proposons d’aligner les FMBs de prédation obtenus à partir de la famille de FOURMI et des familles des lexèmes en (71) pour obtenir le LMB qui est illustré dans la partie supérieure de la Figure 5.7. Les participants de ce LMB sont tous les deux des ANIMAUX : ce qui les distingue à l’intérieur du scénario est le fait que l’un joue le rôle du PRÉDATEUR et l’autre joue le rôle de la PROIE.

La projection du LMB de la Figure 5.7 sur le lexique construit nous permet d’obtenir un nombre très réduit de familles (ou plutôt couples) paradigmatiques, que nous organisons dans la Table 5.3 : dans cette table nous utilisons à nouveau des doubles étiquettes, car le LMB qui produit le paradigme est structuré par des relations argumentales (i.e. PRÉDATEUR et PROIE). La propriété sémantique pertinente pour ce paradigme est le fait que l’élément initial de la famille dénote un animal prédaté par d’autres animaux. D’autre part, nous soulignons que la réalisation morphologique du LMB en Figure 5.7 est très limitée : pour la presque totalité des N_{anim} le nom du prédateur n’appartient pas à la même famille que le nom de sa proie.

En outre, le paradigme en Table 5.3 montre que le nom du PRÉDATEUR, lorsqu’il est présent dans la même famille du N_{anim} qui dénote sa proie, est construit principalement en utilisant le suffixe *-ier*, sauf dans le cas de SERPENTAIRE. Tout au long de ce chapitre, nous allons voir que le suffixe *-ier* réalise une gamme de dérivés extrêmement hétérogènes d’un point de vue sémantique et peut réaliser des fonctions sémantiques très spécifiques.

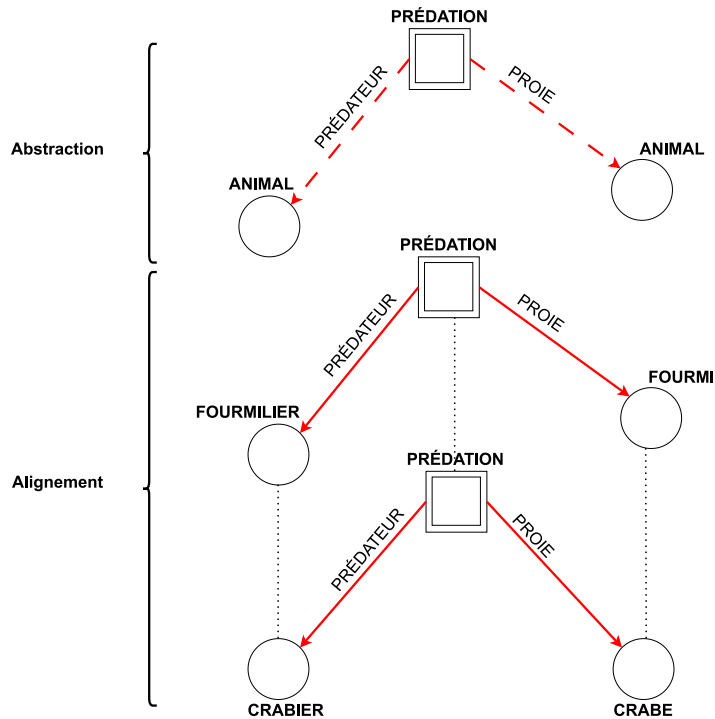


FIGURE 5.7 – Obtention d’un LMB structuré par la relation entre une espèce animale prédatée et une espèce animale prédatrice.

PRÉDATEUR	PROIE
ANIMAL	ANIMAL
FOURMI	FOURMILIER
CRABE	CRABIER
GUÊPE	GUÊPIER
HUÎTRE	HUÎTRIER
SERPENT	SERPENTAIRE

TABLE 5.3 – Paradigme structuré par des événements de prédation entre espèces animales sauvages.

Pour récapituler, la prise en compte de sous-ensembles de familles basées sur N_{anim} en excluant les lexèmes inscrits dans des activités humaines (que nous abordons en Section 5.3) nous a permis d’identifier des LMBs et des paradigmes qui décrivent le comportement animal : les relations sociales entre membres de la même espèce animale sauvage, l’organisation en colonie des insectes et les dynamiques de prédation entre espèces animales. Un premier aspect qui émerge de la prise en compte de ces familles est que parfois les régularités sémantiques qui structurent

les paradigmes peuvent être très spécifiques et le nombre de familles paradigmatiques que l'on peut inscrire dans les paradigmes que nous obtenons est très réduit. Nous allons retrouver cette propriété dans certains paradigmes proposés dans la suite de ce chapitre.

En deuxième lieu, nous avons vu que certaines relations que l'on trouve dans les familles ne peuvent pas être caractérisées au moyen d'actions ou d'activités et ne sont pas de nature argumentale (Voir Sections 5.2.1 et 5.2.2). Dans ces cas, les LMBs que l'on obtient se structurent sous forme de graphes acentriques qui manquent d'un pivot sémantique. Les relations qui structurent ces réseaux acentriques sont des abstractions d'actions de nature variée, comme dans le cas des relations MÂLE-FEMELLE-PETIT chez les animaux sauvages ou la relation INSECTE-NID-COLONIE chez les insectes.

5.3 Animaux et activités humaines

Dans la section précédente, nous avons pris en compte des sous-ensembles de familles dérivationnelles basées sur N_{anim} sauvages en excluant tous les lexèmes relevant d'activités humaines (ex. chasse, pêche, élevage, etc.). Dans cette section, nous étendons l'inventaire des LMBs et paradigmes dérivationnels qui impliquent des N_{anim} en ajoutant ce type de dérivés. Comme nous allons le voir, le lexique construit présente des nombreux dérivés relevant de domaines d'activités dans lesquels les animaux sont des "protagonistes". En outre, la prise en compte des domaines d'activité spécialisés introduit une autre question, à savoir l'accessibilité des lexèmes que l'on trouve dans les familles aux locuteurs. Comme nous allons le voir, la plupart des lexèmes que l'on trouve dans les familles et que l'on peut inscrire dans les paradigmes relevant des activités humaines font partie de lexiques spécialisés et sont très peu attestés ou dans d'autres cas sont marqués comme vieillis par les dictionnaires.

5.3.1 Pêche

Nous commençons en considérant la famille de CREVETTE, dans laquelle on trouve CREVET-TIER. Ce lexème dénote à la fois un bateau utilisé pour la pêche à la crevette et un pêcheur de

crevettes. Sur cette base, nous envisageons que les histoires produites à partir de {CREVETTE.N, CREVETTIER.N} décrivent un scénario de pêche à la crevette. Un exemple d'histoire est fourni en (72) : comme on peut le voir, à part les crevettes et un crevettier dans le sens de 'bateau', cette histoire introduit un équipage de pêcheurs et des haveneaux, qui sont des participants sémantiquement cohérents avec un scénario de pêche. Il est vraisemblable que des PÊCHEURS, un BATEAU DE PÊCHE et des INSTRUMENTS DE PÊCHE soient des participants qui apparaissent de manière régulière dans les histoires produites à partir du couple {CREVETTE.N, CREVETTIER.N} qui décrivent une activité de pêche ou une action qui l'instancie.

- (72) *À bord du crevettier "Crustacé", les pêcheurs ce matin ont pu pêcher une énorme quantité de crevettes. Les nouveaux haveneaux sont beaucoup plus efficaces que les précédents et cela rend la pêche beaucoup plus productive.*

Sur cette base, nous considérons qu'à partir de la famille de CREVETTE, qui contient le couple {CREVETTE.N, CREVETTIER.N}, nous pouvons produire un FMB comme celui en Figure 5.8. Ce LMB contient un crevettier qui joue le rôle du PÊCHEUR, un crevettier qui joue le rôle du BATEAU DE PÊCHE et la crevette qui joue le rôle d'ANIMAL PÊCHÉ. En outre, ce LMB contient également un INSTRUMENT DE PÊCHE de nature artificielle (ex. haveneau, filet, etc.).

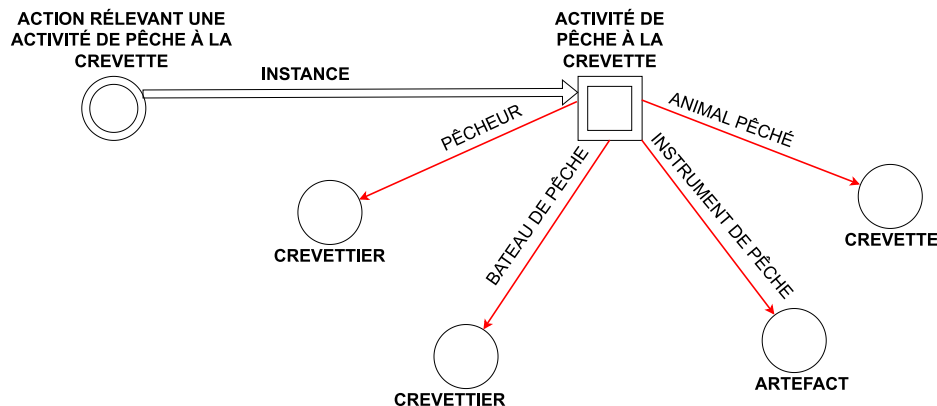


FIGURE 5.8 – FMB décrivant une activité de pêche à la crevette et une action qui l'instancie.

Il se pose la question de trouver d'autres FMBs qui peuvent être alignés avec celui en Figure 5.8. Nous considérons que des FMBs qui contiennent des relations alignables à celles en Figure 5.8 peuvent être constitués à partir d'autres familles dont l'élément initial est un animal pêché. Par exemple, la famille de SARDINE contient SARDINIER, qui dénote un pêcheur de sardines et un bateau utilisé pour la pêche à la sardine, et SARDINAL, qui est un filet utilisé pour pêcher les sardines. Il en va de même pour la famille de HOMARD, qui contient HOMARDIER, le nom du bateau consacré à la pêche du homard. En outre, dans la famille de THON, on trouve THONIER (pêcheur de thons et bateau consacré à la pêche au thon) et THONAIRE, qui est un filet qui sert à prendre les thons. Dans la famille de HARENG, on trouve HARENGUIER, qui est un bateau spécialisé dans la pêche aux harengs, et on trouve également le nom HARENGUEUX qui dénote le même référent. Des exemples d'occurrence de ces deux lexèmes ayant le même sens sont fournis en (73). En outre, dans la famille de HARENG, on a aussi HARENGAISON, qui dénote la pêche aux harengs.

- (73) a. « *Le Duquesne, un remorqueur de la chambre de commerce de Dieppe, rentre au port le Gloire à Marie, un harenguiier fort d'un équipage de vingt-cinq pêcheurs.* » [source *Wiktionnaire*]
- b. « *Nos harengueux repeints en gris, flanqués sur la coque d'un numéro d'immatriculation, deux lettres et deux chiffres, partageaient les bassins avec la deuxième escadre légère, des sous-marins, les torpilleurs de la Dover Patrol.* » [source *Wiktionnaire*]

Par ailleurs, dans la famille de LANGOUSTE le nom LANGOUSTIER dénote à la fois un bateau spécialisé dans la pêche à la langouste et un filet utilisé pour pêcher les langoustes. En dernier lieu, dans la famille de MORUE on a MORUTIER, qui dénote un bateau équipé pour la pêche à la morue et un pêcheur de morue.

Tous ces lexèmes liés à la pêche que nous venons de présenter sont récapitulés dans les sous-ensembles en (74). Nous soulignons que les dérivés que nous venons de mentionner relèvent du domaine spécialisé de la pêche et sont vraisemblablement beaucoup plus accessibles à des experts de pêche qu'à la plupart des locuteurs du français. Cela montre que la prise en compte

de domaines d'activités pour structurer les paradigmes en dérivation peut comporter la difficulté d'inscrire dans des paradigmes des lexèmes qu'un locuteur non spécialiste du domaine ne connaît pas.

- (74) a. {SARDINE.N, SARDINIER.N₁ (pêcheur), SARDINIER.N₂ (bateau), SARDINAL.N (FILET)}
- b. {HOMARD.N, HOMARDIER.N (bateau)}
- c. {THON.N, THONIER.N₁ (bateau), THONIER.N₂ (pêcheur), THONAIRE.N (filet)}
- d. {HARENG.N, HARENGUIER.N (bateau), HARENGUEUX.N (bateau), HARENGAISON.N}
- e. {LANGOUSTE.N, LANGOUSTIER.N₁ (bateau), LANGOUSTIER.N₂ (filet)}
- f. {MORUE.N, MORUTIER.N₁ (bateau), MORUTIER₂ (pêcheur)}

Nous considérons qu'à partir des familles qui contiennent les sous-ensembles en (74) nous pouvons également produire des FMBs décrivant des activités de pêche qui peuvent être alignés avec celui que nous avons proposé en Figure 5.8 pour la pêche à la crevette. Cet alignement nous permet de produire un LMB comme celui qui est proposé en Figure 5.9. Ce LMB représente une situation de pêche qui implique un PÊCHEUR de type ontologique PERSONNE, un BATEAU DE PÊCHE de type ontologique BATEAU, un INSTRUMENT DE PÊCHE de type ARTEFACT (ex. filet, hameçon, etc.) et un ANIMAL PÊCHÉ de type ANIMAL AQUATIQUE.

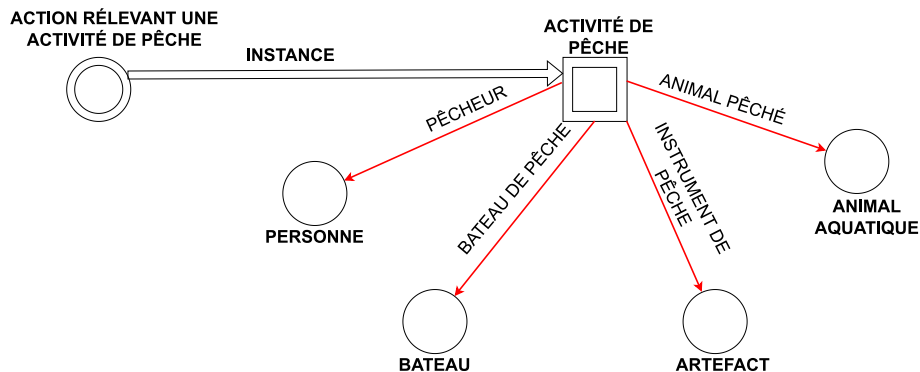


FIGURE 5.9 – LMB décrivant une activité de pêche et l'action qui l'instancie.

La projection du LMB en Figure 5.9 sur le lexique construit permet de produire le paradigme proposé en Table 5.4. Nous rappelons la structure à deux étiquettes qui est adoptée dans les

paradigmes structurés par des activités : la première étiquette est relationnelle (ex. ANIMAL PÊCHÉ) et la deuxième est ontologique (ex. ANIMAL AQUATIQUE). Ce paradigme réunit des familles paradigmatiques basées sur un N_{anim} qui jouent le rôle d'ANIMAL PÊCHÉ dans une activité de pêche. Ce paradigme montre que le nom dénotant le BATEAU DE PÊCHE est régulièrement réalisé dans les familles que nous avons prises en compte et que dans la totalité des cas, il est réalisé au moyen du suffixe *-ier*. En outre, on a un cas où ce même rôle est réalisé par deux dérivés dans la même famille (i.e. le cas de HARENGUIER et HARENGUEUX). Dans ce cas, nous avons deux familles paradigmatiques alignées qui sont obtenues à partir de la même famille de départ (7ème et 8ème ligne de la Table 5.4). Le PÊCHEUR est réalisé de manière moins systématique, mais quand il l'est, il est également réalisé au moyen de suffixation en *-ier*. Pour sa part, l'INSTRUMENT DE PÊCHE est réalisé moins régulièrement dans ces familles. En outre, l'activité est réalisée uniquement dans la famille de HARENG par le nom HARENGAISON. D'autre part, ce nom est principalement associé à la saison où l'on pêche les harengs plutôt qu'à l'activité de pêche en général. Deux exemples d'occurrence de HARENGAISON qui vont dans ce sens sont fournis en (75). Le principal pouvoir prédictif du paradigme en Table 5.4 est donc lié à la réalisation du BATEAU DE PÊCHE dans la famille, qui est régulièrement réalisé au moyen de suffixe *-ier* dans toutes les familles basées sur N_{anim} pêché que nous avons pris en compte.

ANIMAL PÊCHÉ ANIMAL AQUATIQUE	BATEAU DE PÊCHE BATEAU	INSTR. DE PÊCHE ARTEFACT	PÊCHEUR PERSONNE	ACTIVITÉ.N
CREVETTE	CREVETTIER	-	CREVETTIER	-
SARDINE	SARDINIER	SARDINAL	SARDINIER	-
HOMARD	HOMARDIER	-	-	-
THON	THONIER	THONAIRE	THONIER	-
HARENG	HARENGUIER	-	-	HARENGAISON
HARENG	HARENGUEUX	-	-	HARENGAISON
LANGOUSTE	LANGOUSTIER	LANGOUSTIER	-	-
MORUE	MORUTIER	-	MORUTIER	-

TABLE 5.4 – Paradigme dérivationnel contenant des familles basées sur un nom d'ANIMAL PÊCHÉ et qui interviennent dans un scénario de pêche.

- (75) a. *En principe, la saison de pêche ou harengaison, s'ouvre le 1er octobre et se termine vers la mi-décembre*

- b. *En effet, Dan et Stéphanie sont originaires de la région de Dieppe et là-bas, en ce moment, c'est la harengaison, période de l'année où l'on se régale de harengs.*

Une question qui se pose à ce stade est si un scénario de pêche ne peut pas être évoqué également par d'autres familles qui ne contiennent pas de noms d'animal²⁰. Par exemple, nous considérons que des FMBs qui décrivent un scénario de pêche peuvent être également évoqués par des familles dont l'élément initial est un nom qui dénote un INSTRUMENT DE PÊCHE. Les familles en (76) sont basées respectivement sur les noms TRÉMAIL, PALANGRE, LIGNE et CHALUT, qui dénotent tous des entités qui prototypiquement jouent ce rôle dans un scénario de pêche. Dans ces familles, on peut trouver des noms dénominaux qui dénotent un BATEAU DE PÊCHE équipé de cet instrument (TRÉMAILLEUR, PALANGRIER, CHALUTIER, LIGNEUR) et plus rarement le nom du PÊCHEUR qui utilise cet instrument pour pêcher (i.e. ligneur). Dans des familles comme celles de PALANGRE et CHALUT, on trouve également des verbes dénominaux dénotant l'activité de pêche avec l'instrument dénoté par le nom base comme PALANGRER et CHALUTER. Le contenu des familles basées sur un nom dénotant un INSTRUMENT DE PÊCHE est récapitulé en (76).

- (76) a. {TRÉMAIL.N, TRÉMAILLEUR.N (bateau)}
 b. {LIGNE.N, LIGNEUR.N₁(bateau), LIGNEUR.N₂ (pêcheur)}
 c. {PALANGRE.N, PALANGRIER.N (bateau), PALANGRER.V}
 d. {CHALUT.N, CHALUTIER.N (bateau), CHALUTER.V}

Nous considérons qu'à partir des familles en (76) nous pouvons également produire des FMBs décrivant des activités de pêche et que la projection du LMB en Figure 5.9 sur ces familles permet de créer un deuxième paradigme pertinent pour l'activité de pêche, que nous proposons en Table 5.5. Ce qui distingue ce paradigme de celui en Table 5.4 est le fait qu'il contient des familles dont l'élément initial est un INSTRUMENT DE PÊCHE. Comme on peut le voir, les familles dans ce paradigme sont beaucoup moins nombreuses que les familles en Table 5.4 : le participant qui est à nouveau réalisé plus régulièrement est le BATEAU DE PÊCHE

20. Nous avons déjà pu voir qu'un même LMB peut être associé à plusieurs paradigmes dans le Chapitre 4.

qui dans ce paradigme est réalisé au moyen de suffixations en *-ier* et *-eur*. Le paradigme en 5.5 est un autre exemple de paradigme pouvant être rempli par un nombre très limité de familles et qui a un pouvoir prédictif très réduit.

INSTR. DE PÊCHE ARTEFACT	BATEAU DE PÊCHE BATEAU	PÊCHEUR PERSONNE	ACTIVITÉ.V
TRÉMAIL	TRÉMAILLEUR	-	-
PALANGRE	PALANGRIER	-	PALANGRER
CHALUT	CHALUTIER	-	CHALUTER
LIGNE	LIGNEUR	LIGNEUR	-

TABLE 5.5 – Paradigme dérivationnel regroupant des familles basées sur un INSTRUMENT DE PÊCHE.

Pour récapituler, nous venons de proposer deux paradigmes qui regroupent des lexèmes qui s’inscrivent dans une activité de pêche. Le paradigme en Table 5.4 contient des familles dont l’élément initial est un N_{anim} dénotant un ANIMAL PÊCHÉ. Le paradigme en Table 5.5, pour sa part, regroupe des familles dont l’élément initial est un nom qui dénote un INSTRUMENT DE PÊCHE. Ces deux paradigmes ont été obtenus en projetant le même LMB sur les familles, à savoir le LMB en Figure 5.9. Ils contiennent principalement des termes appartenant au lexique spécialisé du domaine de la pêche qui sont peu accessibles aux locuteurs qui n’ont pas de connaissance du domaine : leur principal pouvoir prédictif dérive du fait que le nom du BATEAU DE PÊCHE est systématiquement réalisé dans les familles qui sont organisées dans ces deux paradigmes et que dans la grande majorité des cas il est réalisé au moyen d’une suffixation en *-ier*. Nous allons retrouver un usage régulier de la suffixation en *-ier* pour les noms de bateau en Section 5.5.2, où nous prenons en compte un scénario de transport spécialisé de marchandise. En outre, *-ier* réalise de manière assez régulière des noms dénotant des PÊCHEURS dans le même scénario.

5.3.2 Chasse et animaux “nuisibles”

Nous passons dans cette section au domaine de la chasse. Certaines familles basées sur N_{anim} comme celle de LOUP contiennent des noms qui dénotent le chasseur spécialisé dans la chasse

de l'animal (LOUVETIER) et l'activité de chasse de cet animal (LOUVETERIE). Nous considérons qu'à partir de la famille de LOUP, qui contient le sous-ensemble {LOUP.N, LOUVETIER.N, LOUVETERIE.N}, nous pouvons produire non seulement des histoires qui décrivent le comportement social des loups, mais aussi d'autres histoires qui décrivent des activités de chasse dans laquelle le loup est impliqué en tant qu'espèce animale et joue le rôle d'ANIMAL CHASSÉ. Un exemple d'histoire qui introduit ce scénario est fourni en (77). Dans cette histoire apparaît également la fourrure du loup, que nous pouvons considérer comme une sorte de PRODUIT de l'activité de chasse.

- (77) *Le louwelier, avide de nouvelles fourrures, traqua habilement le loup à travers la forêt.
D'un tir précis avec son fusil, il l'abattit, satisfait.*

En considérant l'histoire en (77) comme étant représentative des histoires que l'on peut produire, nous considérons qu'à partir de la famille de LOUP, on peut produire un FMB comme celui qui est proposé en Figure 5.10 qui décrit une activité de chasse au loup. Ce FMB comprend les rôles suivants : un CHASSEUR de type PERSONNE, une ARME de type ARTEFACT (ex. piège, fusil), un ANIMAL AUXILIAIRE, qui pour la chasse au loup correspond principalement à des chiens et un PRODUIT de l'activité de chasse, qui principalement pour le loup consiste en sa fourrure. Pour ce FMB la distinction mâle-femelle-petit du loup que nous avons abordée en Section 5.2.1 n'est pas pertinente et par conséquent nous considérons que FMB n'est pas interprédictible par rapport à celui qui décrit les interactions sociales entre exemplaires de loup.

Nous considérons que des FMBs similaires à celui que nous venons de proposer et qui décrit la chasse au loup peuvent être produits à partir d'autres familles qui contiennent des lexèmes relevant du domaine de la chasse. Par exemple, dans la famille de CANARD on a CANARDIÈRE, qui dénote un long fusil utilisé pour la chasse aux canards. Dans la famille de RENARD on trouve RENARDIER, qui dénote un chasseur de renards. Dans la famille de BÉCASSE on trouve BÉCASSIER, qui dénote (a) un chasseur de bécasses, (b) plusieurs modèles de fusil utilisés pour la chasse à la bécasse et (c) un chien spécialisé dans la chasse à la bécasse : pour cette raison, nous distinguons BÉCASSIER₁ (chasseur) et BÉCASSIER₂ (fusil) et BÉCASSIER₃ (chien).

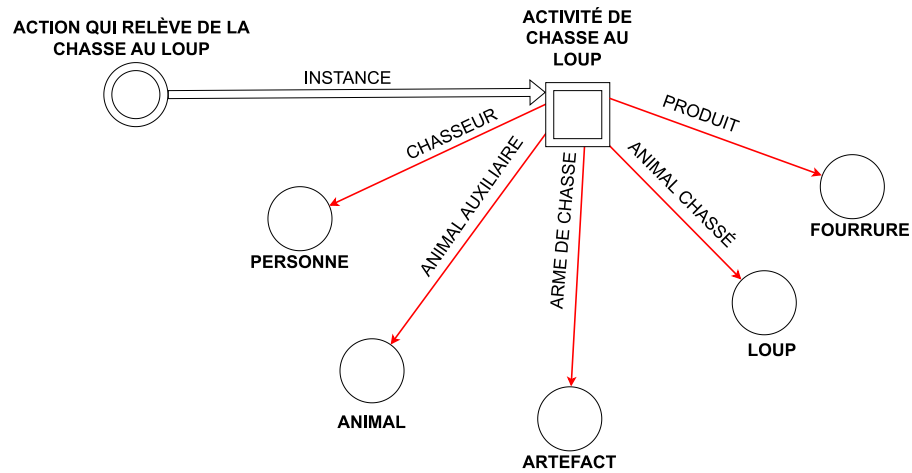


FIGURE 5.10 – FMB décrivant l'activité de chasse au loup et une action qui l'instancie.

En dernier lieu, dans la famille de LIÈVRE on a LEVRIER, qui dénote un chien qui à l'origine était utilisé pour chasser les lièvres. Ces lexèmes sont récapitulés en (78).

- (78) a. {RENARD.N, RENARDIER.N (chasseur)}
- b. {CANARD.N, CANARDIÈRE.N (fusil)}
- c. {BÉCASSE.N, BÉCASSIER.N₁ (chasseur), BÉCASSIER₂ (chien), BÉCASSIER₂ (fusil)}
- d. {LIÈVRE.N, LÉVRIER.N}

Nous considérons qu'à partir des familles des animaux qui sont proposées en (78) nous pouvons également produire des histoires qui décrivent des activités de chasse et donc des FMBs décrivant cette activité. Ces FMBs peuvent être alignés à celui de la Figure 5.10 pour produire le LMB décrivant une activité de chasse que nous proposons en Figure 5.11. Les étiquettes ontologiques que nous utilisons dans ce LMB représentent le dénominateur commun aux activités de chasse des animaux que nous avons mentionnés : l'ANIMAL AUXILIAIRE est typé ANIMAL, car pour la chasse de certains animaux comme le lièvre ou les lapins, on se sert également de furets et de faucons et non seulement de chiens. Une activité de chasse, contrairement à une activité d'élimination d'animal nuisible (que nous abordons dans la suite de cette section), comporte un PRODUIT que nous typons comme PARTIE DU CORPS (ex.

viande de l'animal, fourrure, etc.). Comme pour d'autres LMBs que nous avons abordés, nous pouvons voir que ce LMB comporte des participants qui partagent des propriétés ontologiques, à savoir l'ANIMAL CHASSÉ et l'ANIMAL AUXILIAIRE, qui appartiennent tous les deux à la catégorie ontologique ANIMAL. Ce qui les distingue est leur rôle à l'intérieur du LMB.

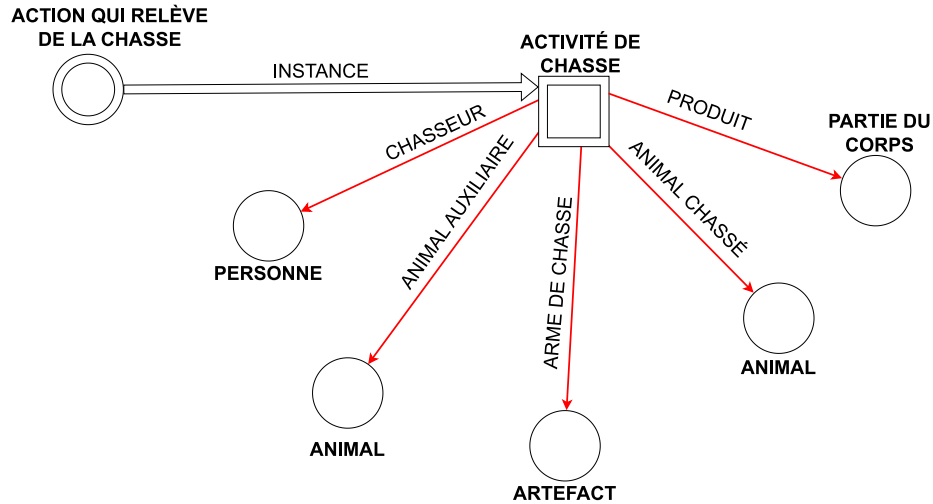


FIGURE 5.11 – LMB décrivant une activité de chasse et l'action qui l'instancie.

La projection du LMB que nous venons de proposer en Figure 5.11 produit le paradigme en Table 5.6. Les familles qui sont organisées dans ce paradigme sont toutes basées sur un N_{anim} qui intervient en tant qu'ANIMAL CHASSÉ dans un scénario de chasse. Comme on peut le voir, elles sont peu nombreuses et contiennent des lexèmes relevant du domaine de la chasse qui sont très peu accessibles pour les locuteurs qui n'ont pas de connaissance du domaine. Le rôle plus fréquemment réalisé dans les familles est celui du CHASSEUR, mais par exemple il n'est pas présent dans les familles de CANARD et LIÈVRE. Le pouvoir prédictif associé au paradigme en Figure 5.6 est donc assez faible, car les familles ne sont pas assez nombreuses et il n'y a pas un rôle qui est systématiquement réalisé de manière morphologique dans les familles, contrairement au BATEAU DE PÊCHE du paradigme précédent en Table 5.5. D'autre part, nous retrouvons à nouveau le suffixe *-ier* qui réalise tous les noms dénotant des CHASSEURS dans la Table 5.6.

En outre, nous soulignons des très rares cas de familles où l'élément initial ne joue pas le rôle d'ANIMAL CHASSÉ mais plutôt d'ANIMAL AUXILIAIRE. Par exemple, dans la famille

ANIMAL CHASSÉ ANIMAL	CHASSEUR PERSONNE	ARME ARTEFACT	ACTIVITE _N	ANIMAL AUXILIAIRE ANIMAL
LOUP	LOUVETIER	-	LOUVETERIE	-
RENARD	RENARDIER	-	-	-
CANARD	CANARDIER	CANARDIÈRE	-	-
BÉCASSE	BÉCASSIER	BÉCASSIER	-	BÉCASSIER
LIÈVRE	-	-	-	LEVRIER

TABLE 5.6 – Paradigmes dérivationnels structurés par une activité de chasse.

de FAUCON on a FAUCONNIER (i.e. chasseur qui chasse à l'aide d'un faucon) et FAUCONNERIE (chasse au faucon) et dans la famille de FURET on a FURETEUR (chasseur au furet) et FURETER (chasser à l'aide d'un furet) : les histoires construites à partir de ces deux familles peuvent également décrire un scénario de chasse comme celui de la Figure 5.11, mais ces deux familles ne rentrent pas dans le paradigme en Table 5.6, car leur élément initial joue un rôle différent dans le LMB décrivant la chasse par rapport aux familles en Table 5.6.

Nous pouvons identifier des relations de nature similaire à celles qui relient les lexèmes dans le paradigme en Table 5.6 en prenant en compte des familles basées sur des N_{anim} considérés comme nuisibles dans certains milieux urbains et non (PIGEON, RAT, SOURIS, TAUPE) et qui sont tués par du personnel spécialisé. Dans ces familles, on trouve de manière régulière des verbes dénotant l'action ou l'activité d'éliminer l'animal qui sont réalisés au moyen de schémas de préfixation en *de-* (ex. DÉRATISER, DÉPIGEONNER, DÉSINSECTISER, DÉTAUPER). En outre, dans ces familles on trouve des noms dénotant la personne chargée de l'élimination de l'animal (DÉRATISEUR, DÉPIGEONNEUR, TAUPIER), des instruments (ou plutôt des armes) dont on se sert pour éliminer l'animal (ex. RATIÈRE, DÉSINSECTISEUR, TAUPIÈRE) et des produits utilisés pour éliminer l'animal (ex. RATICIDE, TAUPICIDE, MOLLUSCICIDE). Ces lexèmes sont récapitulés en (79). Contrairement aux lexèmes que nous venons d'organiser dans le paradigme en Table 5.6, qui relèvent du domaine de la chasse et sont très peu connus, une bonne partie des lexèmes en (79) est largement attestée et accessible à la plupart des locuteurs.

Comme on peut le voir, dans les sous-ensembles en (79) nous ne prenons pas uniquement en compte des lexèmes créés à partir de suffixation, mais aussi ceux qui sont créés au moyen de composition néoclassique dont le composant droit est *-cide*, comme RATICIDE, TAUPICIDE

et MOLLUSCICIDE. Cela ouvre la question de l'intégration dans les familles de départ (et dans les paradigmes) des composés : nous choisissons d'intégrer dans les familles et dans les paradigmes des composés comportant des composants réguliers comme les composés *Xcide* et les composés *Xculteur*, *Xculture*, *Xlogie* et *Xlogue* que nous allons aborder dans les Sections 5.3.3 et 5.3.4. Nous faisons ce choix car nous considérons que les composantes régulières qui forment ces lexèmes (i.e. *-cide*, *-culteur*, *-culture*, *-logie*, *-logue*) ont un comportement proche de celui des affixes, ce qui est en ligne avec la proposition de Lasserre et Montermini (2014b). Plus généralement, la proposition de Lasserre et Montermini (2014a) fait partie des travaux qui considèrent la dérivation affixale et la composition comme des procédés de la même nature qui présentent des différents niveaux de spécification dans les éléments qui les composent (Roché, 2009; Booij, 2010).

- (79) a. {RAT.N, RATIÈRE.N, RATICIDE.N, DÉRATISER.V, DÉRATISEUR.N}
 b. {PIGEON.N, DÉPIGEONNER.V, DÉPIGEONNEUR.N}
 c. {TAUPE.N, TAUPIER.N, TAUPIÈRE.N, TAUPICIDE.N, DÉTAUPER.V}
 d. {SOURIS.N, SOURICIÈRE.N, DÉSORISER.V}
 e. {INSECTE.N, DÉINSECTISER.V, DÉINSECTISEUR.N, INSECTICIDE.N}
 f. {MOUSTIQUE.N, DÉMOUSTIQUER.V}
 g. {MOLLUSQUE.N, MOLLUSCICIDE.N}

Nous considérons qu'à partir des familles qui contiennent les sous-ensembles proposés en (79) on peut produire des histoires qui décrivent l'élimination d'animaux nuisibles. Un exemple d'histoire créée à partir de la famille de RAT est proposé en (80).

- (80) *Le quartier était envahi par les rats, créant une véritable urgence sanitaire. Mme Dupont, excédée, décida d'appeler un dératiseur professionnel. Armé de son expérience et de puissants raticides, l'expert se mit immédiatement au travail pour dératiser le quartier.*

Le LMB proposé en Figure 5.12 représente le contenu régulier des histoires qui décrivent un scénario d'élimination d'animal nuisible : en premier lieu, on a un ÉLIMINATEUR de type

PERSONNE qui se charge d'éliminer un ANIMAL NUISIBLE de type ANIMAL. L'ÉLIMINATEUR peut se servir d'une ARME de type ARTEFACT (ex. piège) ou d'un PRODUIT POUR ÉLIMINER de type SUBSTANCE. Contrairement à l'activité de chasse, nous considérons que l'activité d'élimination d'un animal nuisible n'est pas visée à l'obtention d'un PRODUIT, mais plutôt à l'élimination de l'animal pour des questions sanitaires ou de contrôle de population.

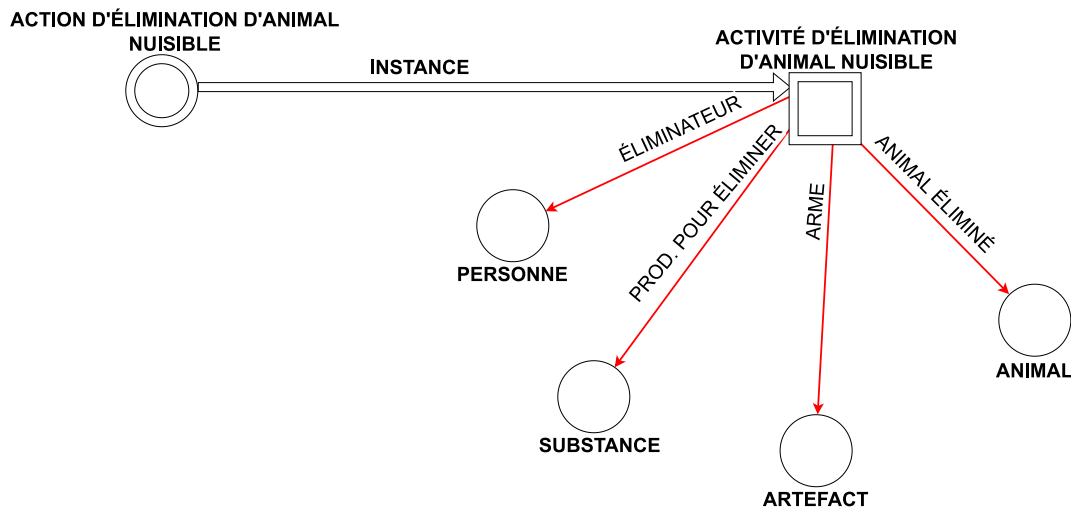


FIGURE 5.12 – LMB décrivant l'activité d'élimination d'animaux nuisibles.

La projection du LMB en Figure 5.12 sur les familles dérivationnelles produit le paradigme qui est proposé en Table 5.7. Comme on peut le voir, le scénario d'élimination d'animal nuisible est caractérisé par une réalisation morphologique dans les familles qui est plus régulière par rapport au scénario de chasse : dans une famille ayant un N_{anim} considéré comme prototypiquement nuisible comme élément initial, on trouve presque systématiquement l'action d'élimination sous forme verbale réalisée au moyen de préfixation en *dé-*, sauf dans le cas de la famille de MOLLUSQUE. Les rôles ÉLIMINATEUR, PRODUIT POUR ÉLIMINER et INSTRUMENT sont réalisés de manière moins systématique dans les familles : lorsque le PRODUIT POUR ÉLIMINER est présent dans la famille, il est réalisé au moyen de composé *Xcide* (RATICIDE, TAUPICIDE, INSECTICIDE, MOLLUSCICIDE). Lorsque l'ARME est présente, elle est réalisée au moyen de suffixation en *-ière* s'il s'agit d'un piège. DÉINSECTISEUR, pour sa part, dénote un appareil électronique.

ANIMAL NUISIBLE ANIMAL	ÉLIMINATION _v	ÉLIMINATEUR PERSONNE	PRODUIT P.ÉLIM. MATIÈRE	ARME ARTEFACT
RAT	DÉRATISER	DÉRATISEUR	RATICIDE	RATIÈRE
TAUPE	DÉTAUPER	TAUPIER	TAUPICIDE	TAUPIÈRE
PIGEON	DÉPIGEONNER	DÉPIGEONNEUR	-	-
SOURIS	DÉROURISER	-	-	SOURICIÈRE
INSECTE	DÉSINSECTISER	-	INSECTICIDE	DÉSINSECTISEUR
MOUSTIQUE	DÉMOUSTIQUER	-	-	-
MOLLUSQUE	-	-	MOLLUSCICIDE	-

TABLE 5.7 – Paradigme dérivationnel qui regroupe des lexèmes inscrits dans une activité d’élimination d’animaux nuisibles.

Pour récapituler, nous avons proposé l’usage de LMBs décrivant des activités de chasse et d’élimination d’animal nuisible pour obtenir des familles paradigmatiques et les organiser dans des paradigmes. Le paradigme obtenu avec la projection du LMB décrivant une activité de chasse contient peu de familles et le peu de lexèmes qu’il contient sont peu fréquents et connus, faisant partie d’un lexique spécialisé. Pour sa part, le paradigme en Table 5.7 a un pouvoir prédictif majeur : la plupart des lexèmes qu’il contient sont accessibles à des locuteurs non experts et il décrit la réalisation systématique de verbes préfixés en *dé-* qui dénotent l’activité d’élimination de l’animal et la réalisation régulière des composés *Xcide* qui dénotent un produit utilisé pour éliminer l’animal nuisible.

5.3.3 Élevage

Nous passons dans cette section au domaine de l’élevage. Nous reprenons la famille de CREVETTE, que nous avons prise en compte dans le cadre d’un scénario de pêche en Section 5.3.1 : cette famille ne contient pas uniquement des lexèmes associés à la pêche, mais aussi des lexèmes qui s’inscrivent dans une activité d’élevage, comme CREVETTICULTEUR (éleveur de crevettes) et CREVETTICULTURE (élevage des crevettes) (81) : ces deux lexèmes relèvent d’un domaine de spécialité d’élevage. De la même manière que POTERIE, le nom CREVETTICULTURE dénote un domaine d’activité : il valide le test proposé par Haas *et al.* (2023, p.62) pour identifier cette catégorie nominale.

(81) {CRÉVETTE.N, CRÉVETTICULTEUR.N, CRÉVETTICULTURE.N}

(82) *Filip a reçu une formation dans le domaine de la crevetticulture.*

Nous considérons qu'à partir de la famille de CREVETTE, dont les lexèmes en (81) font partie, nous pouvons non seulement créer des histoires qui décrivent un scénario de pêche, mais aussi des histoires qui décrivent une activité d'élevage. Un exemple d'histoire de ce type est fourni en (83).

(83) *Le secteur de la crevetticulture est en forte difficulté cette année. Les crevetticulteurs de la région ne peuvent plus assurer la quantité de crevettes qu'ils ont fournie l'année dernière. La sécheresse a endommagé énormément les bassins artificiels et les conditions de vie des crevettes. La production est largement compromise.*

L'histoire en (83) implique des participants que l'on peut considérer comme cohérents avec un scénario d'élevage en milieu aquatique : par exemple, l'histoire parle de bassins artificiels, qui jouent le rôle de LIEU D'ÉLEVAGE (i.e. le lieu où se déroule l'activité).

Nous considérons qu'à partir de la famille de CREVETTE dont le sous-ensemble en (81) fait partie, nous pouvons obtenir un FMB de type activité qui décrit l'activité d'élevage des crevettes. Ce FMB est schématisé en Figure 5.13 : contrairement aux autres *meaning bundles* décrivant des activités que nous avons vues jusqu'ici, il est difficile de considérer que l'on puisse avoir une action délimitée dans le temps et dans l'espace qui relève de l'élevage (contrairement à la fabrication des pots, faire de la chasse ou faire de la pêche de manière ponctuelle). Pour cette raison, le FMB en Figure 5.13 comporte uniquement un volet activité.

Il se pose la question de trouver d'autres FMBs qui peuvent être alignés avec celui de la Figure 5.13. Nous considérons que les scénarios d'élevage pertinents sont ceux qui comportent des élevages en masse, comme les activités réalisées dans le cadre de l'aquaculture ou comme dans l'élevage des escargots ou des vers à soie. D'autre part, nous considérons qu'un élevage en ferme, où les animaux sont élevés et suivis individuellement n'est pas pertinent pour être aligné avec le FMB en Figure 5.13.

Quelles sont donc les familles à partir desquelles on peut produire des FMBs qui peuvent être alignés avec le FMB en Figure 5.13 ? Une première réponse à cette question serait d'aller

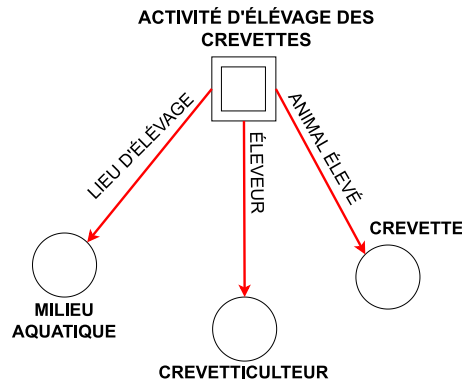


FIGURE 5.13 – FMB décrivant une activité d'élevage de crevettes

voir les familles basées sur des N_{anims} dénotant des animaux qui sont élevés en masse dans des aquacultures ou d'autres types de ferme pour voir lesquelles contiennent des lexèmes qui dénotent des ÉLEVEURS et des LIEUX D'ÉLEVAGE. En revanche, nous constatons une difficulté supplémentaire : la grande majorité des lexèmes qui dénotent une activité d'élevage en masse et des éleveurs qui s'y consacrent sont des composés néoclassiques faisant partie des séries en *Xculteur* et *Xculture* : ces composés comportent la combinaison d'un constituant emprunté au grec ou au latin avec le composant droit *-culture* et *-culteur* (Amiot et Dal, 2008). Ils sont dans la plupart des cas construits sur des thèmes savants qui sont associés aux N_{anims} : par exemple, le nom MYTILICULTURE réfère à l'élevage des moules, SALMONICULTURE réfère à l'élevage des saumons et HÉLICICULTEUR dénote un éleveur d'escargots.

Tout en étant conscients qu'il s'agit d'une approximation, nous considérons que les dérivés qui sont construits sur des thèmes savants sont inscrits dans la même famille du N_{anim} associé à leur thème. Nous faisons ce choix en considérant que les thèmes savants sur lesquels ces composés sont construits font partie de l'espace thématique (Roché, 2010; Roché et Plénat, 2014) des N_{anim} initiaux de la famille. Cela signifie que nous considérons que HÉLICICULTEUR est dans la même famille que ESCARGOT et que MYTILICULTEUR est dans la même famille que MOULE. En (84) nous proposons des familles basées sur N_{anim} qui contiennent des composés *Xculteur* et *Xculture*. En outre, dans les familles de ESTURGEON et ESCARGOT, on trouve également des noms en dénotant le lieu d'élevage (i.e. ESCARGOTIÈRE et ESTURGEONNIÈRE).

- (84) a. {CARPE.N, CARPICULTURE.N, CARPICULTEUR.N}
 b. {TRUITE.N, TRUITICULTEUR.N, TRUITICULTURE.N}
 c. {MOULE.N, MYTILICULTEUR.N, MYTILICULTURE.N}
 d. {SAUMON.N, SALMONICULTEUR.N, SALMONICULTURE.N}
 e. {ESCARGOT.N, HÉLICICULTEUR.N, HÉLICICULTURE.N, ESCARGOTIÈRE.N}
 f. {ESTURGEON.N, ESTURGEONNIÈRE.N }

Si nous considérons les $X_{culteur}$ et $X_{culture}$ comme faisant partie de la même famille que le N_{anim} associé à leur thème savant, à partir des familles en (84) nous pouvons également créer des histoires qui permettent de produire des FMBs qui sont alignables à celui en Figure 5.13. Le LMB proposé en Figure 5.14 décrit toute activité d'élevage d'un animal en masse dans de milieux aquacoles ou dans d'autres types de culture. De la même manière que le FMB d'élevage présenté dans la Figure 5.13, ce LMB comporte uniquement une composante activité.

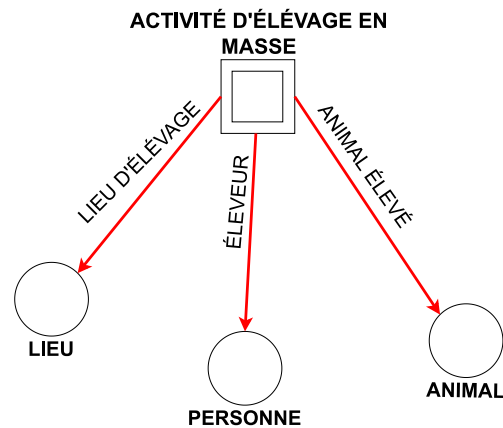


FIGURE 5.14 – LMB décrivant une activité d'élevage en masse d'un animal.

La projection du LMB en Figure 5.14 sur le lexique construit permet d'obtenir les familles paradigmatiques qui sont organisées en Table 5.8. Ce paradigme comporte presque exclusivement des composés $X_{culteur}$ et $X_{culture}$ relevant de domaines spécialisés et construits sur des thèmes savants, mis à part les noms suffixés en *-ière* qui jouent le rôle de LIEU D'ÉLEVAGE. En outre, ces lexèmes relèvent principalement de la terminologie et sont très peu accessibles à des locuteurs non experts. La seule régularité qui caractérise ce paradigme est la coexistence systématique

dans les familles des composés en *Xculteur* et *Xculture* : l'un accompagne toujours l'autre dans la famille. En revanche, cette régularité n'est pas spécifique aux familles basées sur un N_{anim} élevé en masse, car nous retrouvons la même régularité dans les familles basées sur les N_{fru} que nous abordons en Section 5.5 (ex. PRUNICULTEUR-PRUNICULTURE, AGRUMICULTEUR-AGRUMICULTURE) mais elle est plutôt associable à tout type de culture (animale ou végétale).

ANIMAL ELEVÉ ANIMAL	ÉLEVEUR PERSONNE M.	ACTIVITÉ D'ÉLEVAGE ACTIVITÉ	LIEU D'ÉLEVAGE LIEU
CREVETTE	CREVETTICULTEUR	CREVETTICULTURE	-
MOULE	MYTILICULTEUR	MYTILICULTURE	-
SAUMON	SALMONICULTEUR	SALMONICULTURE	-
TRUITE	TRUITICULTEUR	TRUITICULTURE	-
ESTURGEON	-	-	ESTURGEONNIÈRE
ESCARGOT	HÉLICULTEUR	HÉLICULTURE	ESCARGOTIÈRE

TABLE 5.8 – Paradigme dérivationnel structuré par une activité d'élevage en masse. Ce paradigme est caractérisé par la présence régulière des composés néo-classiques *Xculteur* et *Xculture*.

Les familles basées sur N_{anim} élevés dans des fermes comme VACHE ou CHÈVRE, pour leur part, contiennent principalement des dérivés réalisés au moyen d'affixation. On trouve régulièrement des noms dénotant la personne qui s'occupe de l'animal (ex. VACHER, PORCHER, CHEVRIER, MOUTONNIER, ÂNIER, POULARDIER) et des noms dénotant des installations dans la ferme où les animaux sont élevés (ex. VACHERIE, PORCHERIE, CHÈVRERIE, LAPINIÈRE, ASINERIE, POULAILLER).

- (85) a. {VACHE.N, VACHER.N, VACHERIE.N}
 b. {CHÈVRE.N, CHEVRIER.N, CHÈVRERIE.N}
 c. {PORC.N, PORCHER.N, PORCHERIE.N}
 d. {ÂNE.N, ÂNIER.N, ASINERIE.N}
 e. {POULE.N, POULARDIER.N, POULAILLER.N}
 f. {MOUTON.N, MOUTONNIER.N}
 g. {LAPIN.N, LAPINIÈRE.N}

Nous considérons qu'à partir des familles contenant les sous-ensembles en (85) on peut produire des histoires qui décrivent l'activité d'élevage en ferme, comme dans l'histoire en (86) construite à partir de la famille de VACHE.

(86) *Dans la petite ferme, le vacher André s'occupait tendrement de ses vaches. Il connaissait chaque bête par son nom, comprenant leurs moindres meuglements.*

En considérant l'histoire en (86) comme étant illustrative des histoires que l'on peut produire à partir des familles basées sur N_{anim} élevés en ferme, nous proposons le LMB en Figure 5.15 qui décrit une activité d'élevage en ferme. Ce LMB comporte un GARDIEN de type PERSONNE qui s'occupe des animaux, un ANIMAL ÉLEVÉ de type ANIMAL, une INSTALLATION de type ARTEFACT où les animaux sont gardés la plupart du temps et un LIEU D'ÉLEVAGE de type FERME où cette activité a lieu. De la même manière que le LMB décrivant une activité d'élevage en masse des animaux, le LMB en Figure 5.15 ne comporte pas de composante action.

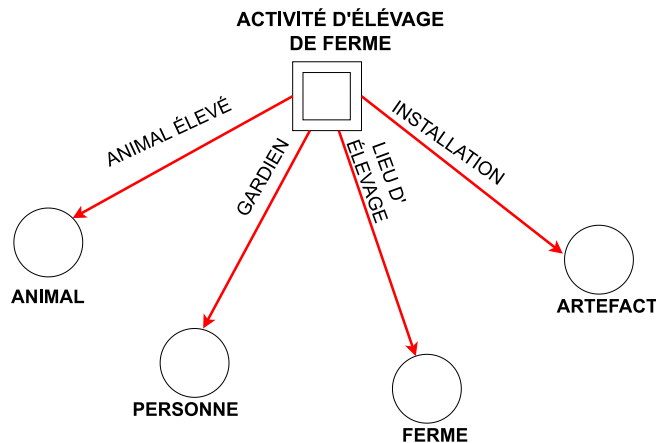


FIGURE 5.15 – LMB décrivant l'activité d'élevage des animaux de ferme.

La projection du LMB en Figure 5.15 sur le lexique construit produit le paradigme dérivationnel proposé en Table 5.9.

Les participants qui sont régulièrement réalisés dans les familles sont le GARDIEN et l'INSTALLATION. Ils sont régulièrement réalisés respectivement au moyen de suffixations en *-ier* et *-erie*. En revanche, une partie des noms contenus dans ce paradigme, comme MOUTONNIER, POULARDIER (qui est construit sur POULARDE, jeune poule engraisée) ou ASINERIE sont très

ANIMAL	GARDIEN PERSONNE M.	INSTALLATION ARTEFACT
VACHE	VACHER	VACHERIE
CHÈVRE	CHEVRIER	CHÈVRERIE
PORC	PORCHER	PORCHERIE
ÂNE	ÂNIER	ASINERIE
POULE	POULARDIER	POULAILLER
LAPIN	-	LAPINIÈRE
MOUTON	MOUTONNIER	-

TABLE 5.9 – Paradigme dérivationnel regroupant des familles paradigmatiques qui interviennent dans une activité d'élevage en ferme.

peu fréquents et marqués par les dictionnaires comme vieilliss, de manière similaire à ce que nous avons vu pour le paradigme relatif à la chasse.

5.3.4 Animaux en tant qu'objet d'étude

Parmi les familles basées sur N_{anim} , on trouve également des familles basées sur des noms dénotant des familles d'animaux comme celle de COLÉOPTÈRE en (87). La famille dérivationnelle de COLÉOPTÈRE contient le composé COLÉOPTEROLOGIE : ce nom fait partie des noms d'activité, car il valide le test en (88). En outre, elle contient plusieurs noms référant à un spécialiste de la coléopterologie, à savoir COLÉOPTERISTE, COLÉOPTEROLOGUE et COLÉOPTEROLOGISTE. Cette concurrence a été décrite par plusieurs études qui se sont intéressées aux réalisations morphologiques des noms de spécialistes (Lignon, 2007; Lasserre et Montermini, 2014a; Villoing et Namer, 2014; Namer et Villoing, 2015).

(87) {COLÉOPTÈRE.N, COLÉOPTÉRISTE.N, COLÉOPTÉROLOGIE.N, COLÉOPTÉROLOGUE.N, COLÉOPTÉROLOGISTE.N}

(88) *Mariame a reçu une formation dans le domaine de la coléopterologie.*

La propriété sémantique du nom COLÉOPTÈRE qui est pertinente pour la présence des autres membres du sous-ensemble en (87) est le fait que les coléoptères, comme pratiquement tous les animaux connus par l'homme, sont étudiés par des experts et font l'objet de disciplines d'étude

scientifique. Nous considérons donc que la propriété sémantique ‘objet d’étude’ est pertinente pour tous les membres de la catégorie ANIMAL connus par l’homme.

Nous considérons que les relations entre les lexèmes en (87) sont structurées par un FMB décrivant une activité d’étude des coléoptères. Les histoires créées à partir de (87) décrivent donc un scénario d’étude de l’animal, avec un SPÉCIALISTE qui a comme OBJET D’ÉTUDE un animal ou une famille d’animaux. D’autres participants récurrents que l’on peut avoir dans les histoires construites à partir de (88) sont des INSTRUMENTS que l’on utilise dans l’étude des coléoptères ou des PRODUITS intellectuels de l’activité d’étude (livres, articles, etc.). Un exemple d’histoire construite sur le sous-ensemble {COLÉOPTÈRE, COLÉOPTERISTE, COLÉOPTEROLOGIE} est proposé en (89). Dans cette histoire, l’article joue le rôle du PRODUIT de l’activité d’étude de Julie.

- (89) *Julie est une coléopteriste experte et elle a récemment écrit un article sur la disparition de plusieurs espèces de coléoptères en France. L’article sera publié sur une revue de coléopterologie.*

Pour trouver d’autres sous-ensembles qui peuvent produire des histoires similaires, nous pouvons chercher parmi les familles de tous les N_{anim} qui dénotent des animaux ou des familles d’espèces animales (ex. insectes, oiseaux, poissons). Des exemples de sous-ensembles de ce type sont fournis en (90) : comme on peut le voir, dans la plupart des cas le nom de la discipline est construit au moyen d’un composé savant comportant *-logie* comme composant droit et un thème savant comme composant gauche (ex. ENTOMOLOGIE, HERPETOLOGIE, ORNITHOLOGIE). La concurrence entre *Xlogie*, *Xlogue* et *Xlogiste* est observable aussi dans la plupart des familles proposées en (90).

Concernant des noms d’INSTRUMENT utilisés pour l’étude de l’animal qui sont inscrits dans le LMB, nous observons qu’ils ne sont presque jamais réalisés dans les familles : dans la famille de INSECTE on a INSECTIER, qui dénote un casier utilisé pour conserver les insectes, mais il n’est pas sûr qu’un spécialiste d’insectes se serve d’un insectier. Pour citer deux autres instruments, dans la famille de OISEAU on trouve ORNITHOSCOPE, qui dénote une longue vue

utilisée pour observer les oiseaux²¹ et dans la famille de INSECTE on a INSECTOSCOPE, qui dénote un ancien modèle de microscope qui était utilisé autrefois pour étudier l'anatomie des insectes. En revanche, ces deux lexèmes ne sont pas pertinents pour la structure paradigmatique de ces familles car ils sont tous les deux extrêmement rares.

- (90) a. {INSECTE.N, ENTOMOLOGIE.N, ENTOMOLOGISTE.N, INSECTOLOGUE.N, INSECTOLOGIE.N, INSECTOSCOPE.N}
- b. {ARAIGNÉE.N, ARANÉOLOGIE.N, ARANÉOLOGUE.N, ARANÉOLOGISTE.N}
- c. {REPTILE.N, HERPETOLOGIE.N, HERPETOLOGUE.N, HERPETOLOGISTE}
- d. {OISEAU.N, ORNITHOLOGIE.N, ORNITHOLOGUE.N, ORNITHOLOGISTE.N, ORNITHOSCOPE.N}
- e. {PRIMATE.N, PRIMATOLOGUE.N, PRIMATOLOGIE.N, PRIMATOLOGISTE.N}
- f. {POISSON.N, ICTHYOLOGIE.N, ICTHYOLOGUE.N, ICTHYOLOGISTE.N}

Nous considérons qu'à partir de familles qui contiennent les sous-ensembles en (90) et à partir de la famille de COLÉOPTÈRE dont nous sommes partis on peut produire de manière régulière des histoires qui décrivent une activité d'étude d'un animal. Le contenu régulier de ces histoires est représenté dans le LMB en Figure 5.16.

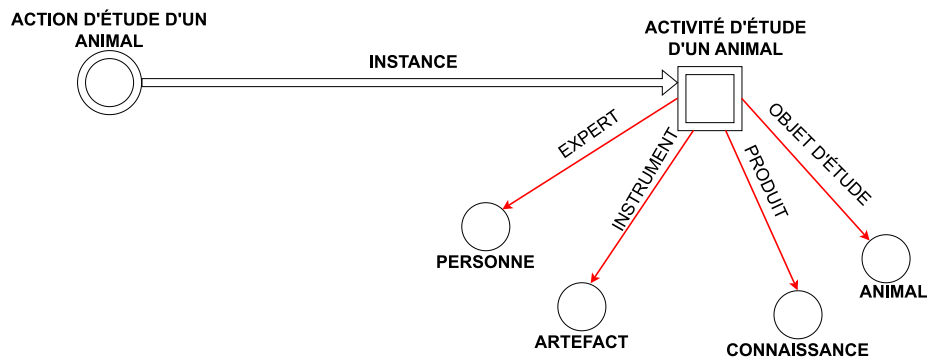


FIGURE 5.16 – LMB décrivant l'activité d'étude d'un animal. Ici l'étiquette ANIMAL attribuée au participant OBJET D'ÉTUDE se situe sur tous les niveaux de classification animale (espèces, familles d'animaux, etc.)

21. <https://ornithoscopie.monsite-orange.fr/>

La projection du LMB en Figure 5.16 sur les familles contenues dans le lexique construit nous permet d'obtenir le paradigme dérivationnel en Table 5.10. Ce qui caractérise ce paradigme est le fait qu'à partir de la même famille on peut produire plusieurs familles paradigmatiques alignées, suite au phénomène de concurrence cité auparavant. En outre, on a de la concurrence au niveau des affixes, mais aussi au niveau du thème : par exemple, un ENTOMOLOGISTE et un INSECTOLOGUE étudient tous les deux les insectes. Sans compter la présence épisodique des composés INSECTOSCOPE et ORNITHOSCOPE, au niveau des trois premières colonnes le pouvoir prédictif du paradigme en Table 5.10 est assez faible pour plusieurs raisons : la présence systématique de composés savants relevant de domaines spécialisés, le fait que les composés savants sont construits dans la plupart des cas sur un thème suppletif associé à l'animal et la concurrence au niveau des thèmes et des affixes.

ANIMAL ÉTUDIÉ ANIMAL	SPÉCIALISTE PERSONNE	ACTIVITÉ D'ÉTUDE	INSTRUMENT ARTEFACT
COLÉOPTÈRE	COLÉOPTERISTE	COLÉOPTEROLOGIE	-
COLÉOPTÈRE	COLÉOPTEROLOGISTE	COLÉOPTEROLOGIE	-
REPTILE	HERPÉTOLOGUE	HERPETOLOGIE	-
REPTILE	HERPÉTOLOGISTE	HERPETOLOGISTE	-
INSECTE	ENTOMOLOGISTE	ENTOMOLOGIE	INSECTOSCOPE
INSECTE	INSECTOLOGUE	INSECTOLOGIE	INSECTOSCOPE
ARAIGNÉE	ARENÉOLOGISTE	ARANÉOLOGIE	-
ARAIGNÉE	ARANÉOLOGUE	ARANÉOLOGIE	-
PRIMATE	PRIMATOLOGUE	PRIMATOLOGIE	-
PRIMATE	PRIMATOLOGISTE	PRIMATOLOGIE	-
OISEAU	ORNITHOLOGUE	ORNITHOLOGIE	ORNITHOSCOPE
OISEAU	ORNITHOLOGISTE	ORNITHOLOGIE	ORNITHOSCOPE
POISSON	ICHTYOLOGISTE	ICHTYOLOGIE	-
POISSON	ICHTYOLOGUE	ICHTYOLOGIE	-

TABLE 5.10 – Paradigme dérivationnel regroupant des lexèmes intervenant dans une activité d'étude d'un animal.

5.4 Bilan intermédiaire

La prise en compte des familles basées sur N_{anim} nous a permis d'identifier une variété de paradigmes dérivationnels obtenus à partir de LMBs décrivant des scénarios divers. En Section 5.2 nous avons produit des paradigmes obtenus à partir de scénarios décrivant le comportement

animal : l'organisation sociale des exemplaires d'une espèce, les insectes et leur organisation en colonie et les relations de prédation entre espèces animales sauvages. La particularité des LMBs décrivant les relations sociales entre animaux sauvages et la vie des insectes en colonie est le fait que les relations contenues dans ces *meaning bundles* sont des abstractions d'événements de nature diverse. Ces deux LMBs décrivent des régularités sémantiques très spécifiques qui ne concernent qu'une partie du royaume animal. Les paradigmes que nous avons produits à partir de ces LMBs sont assez petits en termes d'extension et du nombre de familles qui peuvent y rentrer : cela vaut surtout pour le paradigme décrivant les relations entre les insectes, leur colonie et leur nid, qui inclut très peu de familles.

D'autre part, la prise en compte de domaines de spécialité relevant des activités humaines où les animaux sont des protagonistes a révélé d'autres difficultés dans l'identification des paradigmes : certains paradigmes, comme celui qui s'inscrit dans une activité de pêche, permettent de décrire des régularités comme la présence régulière des noms de BATEAU DE PÊCHE en *-ier*, mais les lexèmes qui y sont inscrits sont très peu connus aux locuteurs non experts. Le paradigme inscrit dans une activité de chasse, pour sa part, comporte très peu de familles qui contiennent des lexèmes très peu connus. Le paradigme qui contient des lexèmes plus accessibles à des locuteurs non experts est celui qui s'inscrit dans le cadre de l'élimination d'animaux nuisibles et qui permet de décrire la formation régulière des verbes en *dé-* qui dénotent le procédé d'élimination de l'animal nuisible.

Les paradigmes relevant de l'élevage en masse d'animaux et l'activité d'étude de l'animal, pour leur part, comportent deux difficultés principales : ils contiennent principalement des composés savants construits sur des thèmes supplétifs et leur contenu relève principalement de la terminologie qui est peu accessible aux locuteurs non experts. Le paradigme inscrit dans le cadre d'une activité d'élevage en ferme, pour sa part, contient lui aussi des lexèmes très peu fréquents et accessibles aux locuteurs.

Pour récapituler, lorsque nous nous plaçons au niveau d'activités humaines spécifiques et douées d'une propre terminologie, nous nous retrouvons avec des paradigmes contenant principalement des lexèmes très peu attestés et prédictibles, sauf des rares exceptions comme le cas

des verbes dénominaux en *dé-* dénotant une action d'élimination d'animal nuisible.

5.5 Fruits

Dans cette section, nous nous intéressons à des familles contenant des noms de fruits (N_{fru}) pour voir quels paradigmes nous pouvons obtenir en partant par cette catégorie ontologique. Nous partons de la proposition faite par Fradin (2020a), qui considère que les paradigmes que l'on peut produire à partir des N_{fru} sont entièrement basés sur les propriétés ontologiques attribuables aux fruits, à savoir le fait qu'ils poussent sur des arbres fruitiers et que ces arbres fruitiers poussent dans des terrains, (voir Section 2.4.1 du Chapitre 2).

5.5.1 Production fruiticole

Pour commencer, considérons les lexèmes contenus dans la famille de POMME et de PRUNE en (91). Dans ces familles, nous trouvons à la fois des éléments formés par suffixation et des composés en *Xculteur* et *Xculture* que nous avons déjà pu voir dans le cadre de l'élevage et que nous avons intégrés aux familles et aux paradigmes.

- (91) a. {POMME.N, POMMIER.N, POMMERAIE.N, POMMICULTEUR.N, POMMICULTURE.N}
 b. {PRUNE.N, PRUNIER.N, PRUNERAIE.N, PRUNICULTEUR.N, PRUNICULTURE.N}

Contrairement à Fradin (2020a), nous considérons que les lexèmes contenus dans les familles en (91) ne dépendent pas exclusivement des propriétés naturelles des pommes et des prunes, mais doivent plutôt être inscrits dans un scénario dénotant une activité humaine basée sur la production de ces fruits. Désormais, la presque totalité des fruits consommés par l'homme est le résultat d'activités de production fruiticole et non de procédés complètement naturels. L'homme produit une grande partie des fruits qu'il consomme et le fait d'être cultivé est désormais une propriété constitutive des plantes fruitières et des fruits que ces plantes produisent. De ce fait, les fruits sont semblables à des artefacts : les agriculteurs cultivent les plantes fruitières en ayant

comme objectif la production de fruits avec certaines caractéristiques physiques et de goût avec la finalité d'être consommés.

Pour donner une idée du genre d'histoires que l'on peut produire à partir de familles comme celles en (91), nous proposons en (92) un exemple d'histoire qui contient en même temps tous les lexèmes contenus dans la famille de PRUNE en (91b). Cette histoire montre que les sens associés aux lexèmes en (91b) peuvent tous être inscrits dans un scénario de production fruiticole, qui correspond à l'origine prototypique des prunes.

- (92) *Le pruniculteur dévoué travaillait avec passion au cœur de sa pruneraie, entouré de majestueux pruniers. Avec expertise, il pratiquait l'art de la pruniculture, donnant vie à de délicieuses prunes. Chaque saison, les arbres se couvraient d'une profusion de fleurs blanches, annonçant la promesse de récoltes abondantes. Grâce au savoir-faire du pruniculteur, la pruneraie prospérait.*

En considérant l'histoire en (92) comme étant illustrative de la plupart des histoires produites à partir de la famille de PRUNE, nous proposons en Figure 5.17 un FMB qui décrit l'activité de production des prunes. Dans ce FMB, nous considérons que le prunier joue le rôle du MOYEN, car il subit un changement dans le processus de production du fruit. Par rapport à cette activité, la prune est le PRODUIT. En outre, on a un AGRICULTEUR de type PERSONNE et des INSTRUMENTS éventuels de type ARTEFACT dont l'agriculteur peut se servir dans la cultivation.

Nous considérons que les familles à partir desquelles on peut produire des histoires proches à celle en (92) et des FMBs qui peuvent être alignés avec celui en Figure 5.17 sont des familles dont l'élément initial est un fruit cultivé qui intervient en tant que PRODUIT dans une activité de production fruiticole. Les familles en (93) peuvent également produire des histoires qui décrivent des activités de production fruiticole qui impliquent un AGRICULTEUR, un ou plusieurs INSTRUMENTS utilisés dans la cultivation, etc. Comme nous l'avons déjà vu avec d'autres domaines d'activité, les familles en (93) contiennent également des dérivés qui font

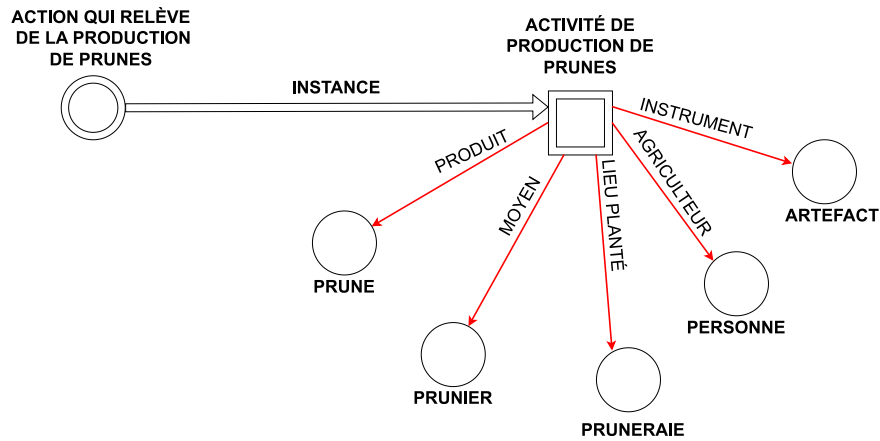


FIGURE 5.17 – FMB décrivant la production des prunes.

partie de lexiques spécialisés : par exemple, un MELONNIER est producteur de melons et la FIGUICULTURE est l'activité de production des figes.

- (93) a. {FIGUE.N, FIGUIER.N, FIGUERAIE.N, FIGUICULTEUR.N, FIGUICULTURE.N}
 b. {BANANE.N, BANANIER.N, BANANERAIE.N, BANANICULTURE.N}
 c. {PÊCHE.N, PÊCHER.N, PÊCHERAIE.N}
 d. {MELON.N, MELONNIER.N, MELONNIÈRE.N}

Le LMB que nous proposons en Figure 5.18 généralise le contenu des FMBs que l'on peut produire à partir des familles en (93) et de la famille de PRUNE et décrit de manière générale une activité de production fruiticole. Nous retrouvons les mêmes participants que nous avons introduits avec le FMB en Figure 5.17, la différence est sur le plan de la généralisation ontologique (ex. le PRODUIT est un FRUIT).

Le paradigme proposé en Table 5.11 est le résultat de la projection du LMB en Figure 5.18 sur le lexique construit. Ce paradigme montre que dans les familles basées sur N_{fru} le MOYEN (la PLANTE) et le LIEU PLANTÉ sont systématiquement réalisés de manière morphologique dans les familles et avec des moyens formels réguliers. Pour les noms de la plante, nous avons presque systématiquement une suffixation en *-ier* qui intervient, sauf dans la famille de MELON où la PLANTE a le même nom du fruit et le suffixe *-ier* est pris par l'AGRICULTEUR. En outre, le nom du LIEU PLANTÉ est régulièrement réalisé au moyen du suffixe *-eraie*, sauf dans

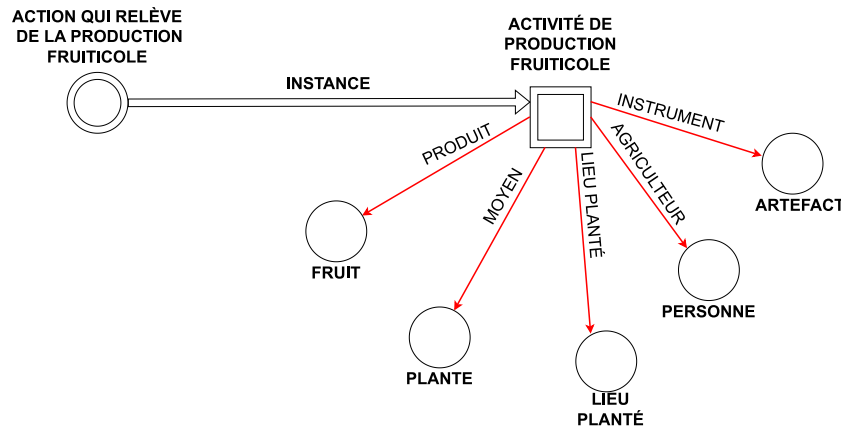


FIGURE 5.18 – LMB décrivant une activité de production fruiticole et l'action qui l'instancie.

PRODUIT FRUIT	MOYEN PLANTE	LIEU PLANTÉ TERRAIN	AGRICULTEUR PERSONNE	ACTIVITÉ PROD. FRUIT.
PRUNE	PRUNIER	PRUNERAIE	PRUNICULTEUR	PRUNICULTURE
POMME	POMMIER	POMMERAIE	POMMICULTEUR	POMMICULTURE
FIGUE	FIGUIER	FIGUERAIE	FIGUICULTEUR	FIGUICULTURE
BANANE	BANANIER	BANANERAIE	-	BANANICULTURE
MELON	MELON	MELONNIÈRE	MELONNIER	-
PÊCHE	PÊCHER	PÊCHERAIE	-	-

TABLE 5.11 – Paradigme dérivationnel structuré par une activité de production fruiticole.

le cas de MELONNIÈRE. Lorsque les noms de l'activité fruiticole et de l'AGRICULTEUR sont présents dans la famille, ils sont dans la plupart des cas réalisés au moyen de composés *Xculteur* et *Xculture*, (sauf pour MELONNIER et dans d'autres cas du même type comme TOMATIER, qui dénote un cultivateur de tomates.).

Pour récapituler, le paradigme en Table 5.11 n'est pas structuré par les propriétés ontologiques du fruit mais plutôt sur son rôle en tant que PRODUIT dans le cadre d'une activité de production fruiticole. Ce paradigme présente des régularités principalement au niveau de la réalisation du MOYEN et du LIEU PLANTÉ.

5.5.2 Transport spécialisé

Parmi les lexèmes que nous avons inscrits dans le paradigme structuré par une activité de production fruiticole, on trouve BANANIER, qui dénote un arbre à bananes. Comme nous l'avons anticipé dans l'introduction de ce chapitre, BANANIER ne réfère pas uniquement à un arbre à bananes, mais aussi à un cargo utilisé pour le transport des bananes et à un gâteau à base de banane. Cela signifie que les histoires produites à partir du couple $\{\text{BANANE.N}, \text{BANANIER.N}\}$ et plus généralement par la famille de BANANE ne vont pas uniquement décrire des activités fruiticoles mais impliqueront également des cargos et des gâteaux. Dans la famille de BANANE, nous distinguons donc BANANIER₁ (arbre), BANANIER₂ (cargo) et BANANIER₃ (gâteau). Après avoir identifié le LMB de production fruiticole en Section 5.5.1, nous savons déjà comment décrire la relation sémantique qui relie BANANIER₁ à BANANE et dans quel scénario l'inscrire : dans cette section et dans la suivante nous traitons BANANIER₂ et BANANIER₃.

Nous considérons que les histoires construites à partir de $\{\text{BANANE.N}, \text{BANANIER.N}_2\}$ décrivent un scénario de transport de bananes, où le bananier joue le rôle du navire qui transporte une cargaison de bananes d'un lieu de récolte ou de production à un lieu de distribution. Un exemple d'histoire que l'on peut construire sur ce couple est fourni en (94). L'histoire en (94) évoque des participants sémantiquement cohérents avec un scénario de transport maritime, comme des marins ou les marchés locaux, qui jouent le rôle de DESTINATION du transport. En outre, cette histoire ne fait aucune référence à la production de bananes et n'implique pas des bananeraies ou des agriculteurs producteurs de bananes. Le scénario de production des bananes est donc un scénario distinct de celui qui décrit leur transport à bord de bananiers. Pour le dire autrement, un bananier (cargo) n'est pas interprétable avec bananier (plante).

- (94) *Le bananier naviguait sur les eaux, transportant une cargaison de bananes fraîches. Les marins s'assuraient que les fruits restaient en bon état pendant le voyage afin qu'ils puissent être distribués et appréciés par les consommateurs. Une fois arrivées à destination, les bananes étaient déchargées avec soin, prêtes à être vendues dans les marchés locaux.*

En considérant l'histoire en (94) comme étant représentative des histoires que l'on peut créer

à partir de $\{\text{BANANE.N}, \text{BANANIER}_2\}$, en Figure 5.19 nous proposons un FMB obtenu à partir de la famille de BANANE qui décrit le transport des bananes. En premier lieu, la banane joue le rôle du BIEN TRANSPORTÉ et le bananier joue le rôle du MOYEN DE TRANSPORT. En outre, nous incluons dans ce FMB l'ÉQUIPAGE du bananier, la SOURCE et la DESTINATION du déplacement. Ces trois derniers rôles ne sont pas réalisés dans la famille de BANANE, mais sont sémantiquement cohérents avec un scénario de transport maritime spécialisé : si on transporte une cargaison de bananes, il doit y avoir un groupe de personnes à bord du bananier et il doit y avoir un lieu de départ (SOURCE) et un lieu d'arrivée (DESTINATION).

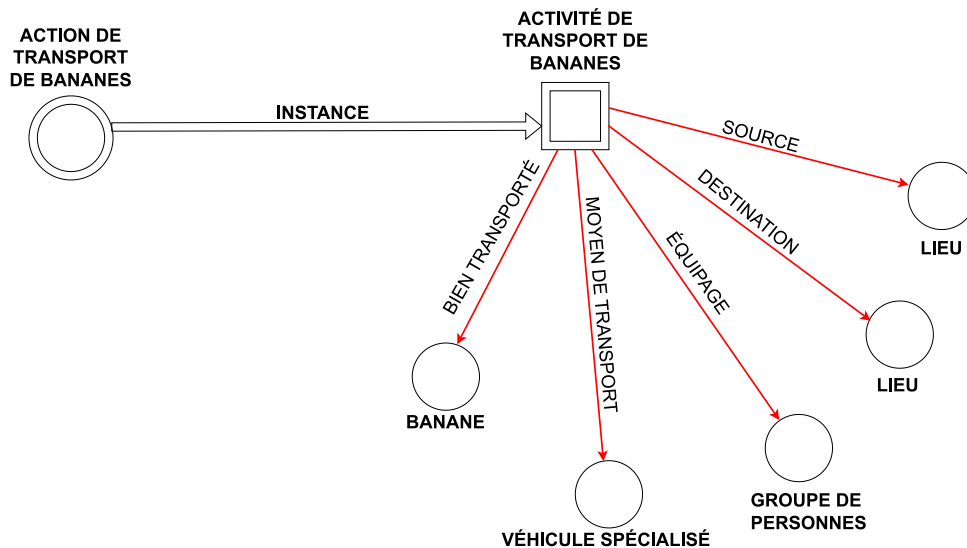


FIGURE 5.19 – FMB décrivant une activité de transport de bananes dans un cadre de distribution commerciale produit à partir de la famille de BANANE.

Il reste à définir quelle est la propriété sémantique de BANANE qui est pertinente pour la relation avec BANANIER₂ et quels autres FMBs peuvent être alignés avec celui en Figure 5.19. Si la propriété pertinente pour cette relation était la nature fruit de BANANE, on s'attendrait à avoir dans beaucoup de familles basées sur N_{fru} le nom du cargo spécialisé dans leur transport. Par exemple, dans la famille de ANANAS, MANGUE, AVOCAT ou PAPAYE, on devrait avoir des noms dérivés dénotant un cargo ou un moyen de transport spécialisé dans leur transport. Cependant, ces dérivés ne sont pas présents dans les familles de ces fruits. La nature de fruit donc n'est pas pertinente pour la relation entre BANANE et BANANIER₂.

Une deuxième possibilité est de considérer que c'est le fait que les bananes sont commercialisées qui est pertinent pour la relation entre BANANE et BANANIER₂. Ou encore plus précisément, nous pouvons considérer que c'est le fait que les bananes sont des marchandises qui ont besoin d'être transportées dans des véhicules spécialisés et que pour cette raison on a des bateaux qui ne transportent que des bananes et que l'on appelle ces bateaux BANANIER. La spécialisation du BANANIER consiste à être équipé pour pouvoir transporter des bananes dans des conditions de température qui permettent de les conserver tout au long du voyage.

Selon cette deuxième hypothèse, pour pouvoir identifier d'autres FMBs similaires à celui qui est proposé en Figure 5.19 nous devons prendre en compte des familles basées sur des noms qui dénotent des marchandises qui nécessitent un transport spécialisé avec des navires qui sont équipés *ad-hoc*. Des exemples de familles partielles pertinentes dans ce sens sont présentés en (95) : il s'agit de couples contenant un nom dénotant une MARCHANDISE SPÉCIALE (ex. PÉTROLE, CÉRÉALE, MÉTHANE, GAZ) et un nom dénominal dénotant le véhicule spécialisé pour leur transport (ex. PÉTROLIER, CÉRÉALIER, MÉTHANIER, GAZIER). Nous précisons que tous ces noms de véhicule dénotent des bateaux. Nous retrouvons donc une série de noms de bateaux spécialisés en *-ier* après celle que nous avons trouvée en Section 5.3.1 pour le domaine de la pêche.

Sur cette base, nous considérons que la propriété sémantique 'marchandise spéciale' est la propriété sémantique de BANANE qui est pertinente pour la relation avec BANANIER₂. La série des noms de bateaux spécialisés en (*-ier*) que nous proposons en (95) est construite sur une variété de noms dont la propriété sémantique commune est le fait d'être une marchandise avec des besoins de transports spéciaux (température, espace de stockage, etc.).

- (95) a. {BANANE.N, BANANIER.N₂}
- b. {PÉTROLE.N, PÉTROLIER.N}
- c. {CÉRÉALE.N, CÉRÉALIER.N}
- d. {PHOSPHATE.N, PHOSPHATIER.N}
- e. {MÉTHANE.N, MÉTHANIER.N}

- f. {ASPHALTE.N, ASPHALTIER.N}
- g. {HUILE.N, HUILIER.N}
- h. {GAZ.N, GAZIER.N}
- i. {BUTANE.N, BUTANIER.N}
- j. {VOITURE.N, VOITURIER.N}²²

Nous considérons que les histoires produites à partir des familles dont les couples en (95) font partie auront un contenu sémantique régulier et décriront une activité de transport spécialisé de marchandise d'un lieu de production à un lieu de distribution à bord d'un navire spécialisé. Ce contenu sémantique commun est représenté dans le LMB en Figure 5.20. Ce LMB comporte un BIEN TRANSPORTÉ de type MARCHANDISE, un MOYEN DE TRANSPORT de type NAVIRE SPÉCIALISÉ, un ÉQUIPAGE de type GROUPE DE PERSONNES, et une SOURCE et une DESTINATION de type LIEU. La projection du LMB en Figure 5.20 sur les familles donne le paradigme dérivationnel binaire en Table 5.12, car les autres rôles de ce LMB ne sont pas réalisés dans les familles en question.

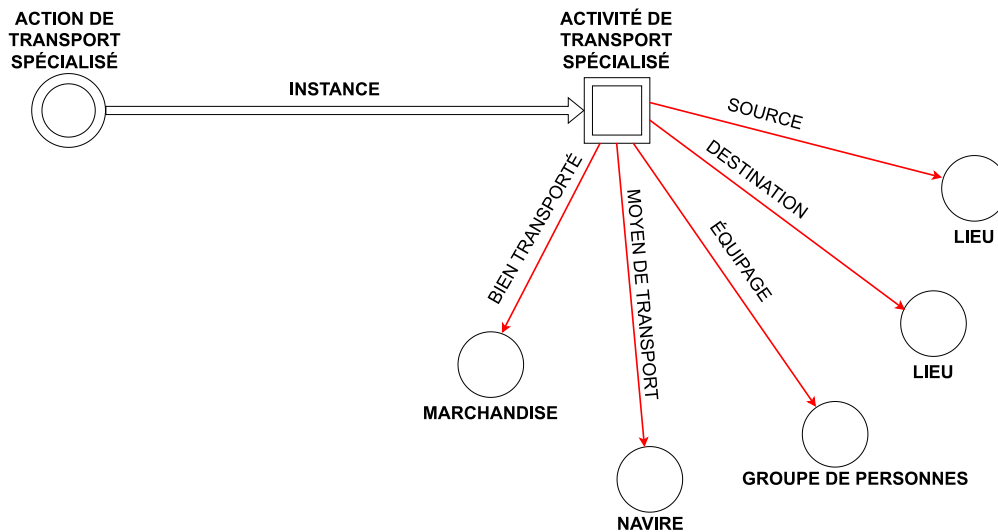


FIGURE 5.20 – LMB décrivant une activité de transport d'un bien commercialisé par un véhicule spécialisé.

22. VOITURIER dénote à la fois une personne chargée de s'occuper de la voiture d'un client et un navire spécialisé dans le transport de véhicules.

BIEN TRANSPORTÉ MARCHANDISE	MOYEN DE TRANSPORT VÉHICULE SPÉCIALISÉ
BANANE	BANANIER
PÉTROLE	PÉTROLIER
GAZ	GAZIER
ASPHALTE	ASPHALTIER
CÉRÉALE	CÉRÉALIER
BUTANE	BUTANIER
VOITURE	VOITURIER

TABLE 5.12 – Paradigme structuré par une activité de transport spécialisé d’une marchandise

Sur le plan prédictif, le paradigme en Table 5.12 montre la réalisation régulière dans les familles basées sur un nom dénotant des marchandises spéciales du nom de bateau spécialisé en *-ier*. En outre, ce paradigme montre que les propriétés sémantiques pertinentes pour la structuration des paradigmes ne sont pas forcément toujours les plus saillantes que l’on peut associer à un lexème : BANANE a bien sa place dans ce paradigme malgré le fait que la propriété sémantique ‘marchandise spéciale’ n’est pas la propriété sémantique principale qu’on pourrait lui associer. Plus généralement, le paradigme en Table 5.12 est un exemple de paradigme dérivationnel décrivant une relation sémantique très spécifique et qui est régulièrement réalisée dans les familles au moyen du même suffixe, à savoir *-ier*.

5.5.3 Cuisine et recettes

Nous avons annoncé précédemment que la famille de BANANE inclut également un lexème BANANIER₃ dénotant un gâteau préparé avec des bananes. Nous considérons que la propriété sémantique de BANANE qui est sélectionnée par cette relation n’est pas sa nature de fruit, mais plutôt le fait qu’il s’agit d’un fruit comestible qui peut être utilisé comme ingrédient dans des préparations. Pour cette raison, nous envisageons que les histoires produites à partir du couple {BANANE, BANANIER₃} décrivent un scénario de préparation d’un bananier, avec un CUISINIER qui le prépare, des INGRÉDIENTS et des INSTRUMENTS DE CUISINE utilisés pour la préparation (ex. four, ustensiles de cuisine, etc.). Un exemple d’histoire qui réalise ce scénario est fourni en (96) : cette histoire n’a rien à voir ni avec le scénario de production des

bananes que nous avons décrit en Section 5.5.1 ni avec le scénario de transport des bananes que nous avons décrit en Section 5.5.2. Sur cette base, nous considérons qu'un bananier (gâteau) n'est pas interprédictible par rapport à la plante ou au navire spécialisé dans le transport.

- (96) *Julie a préparé un bananier pour le repas de dimanche avec des bananes qui lui restaient et du chocolat pâtissier. La préparation lui a pris environ 15 minutes, puis elle a mis le bananier au four.*

Sur cette base, nous considérons qu'à partir de la famille de BANANE, qui à son tour contient le sous-ensemble {BANANE.N, BANANIER.N₃}, on peut produire le FMB qui est proposé en Figure 5.21. Ce FMB décrit une activité de préparation d'un bananier effectuée par une personne qui joue le rôle du CUISINIER en utilisant plusieurs INGRÉDIENTS de type matière comestible (ex. banane, chocolat, œufs, sucre) et des OUTILS DE CUISINE de type ARTEFACT (ex. mixeur, four). Nous précisons que dans ces histoires plusieurs ingrédients et plusieurs outils de cuisine interviennent en même temps. Comme pour d'autres réseaux de type activité que nous avons présentés, nous considérons que cette activité peut être instanciée par une action délimitée dans le temps et dans l'espace.

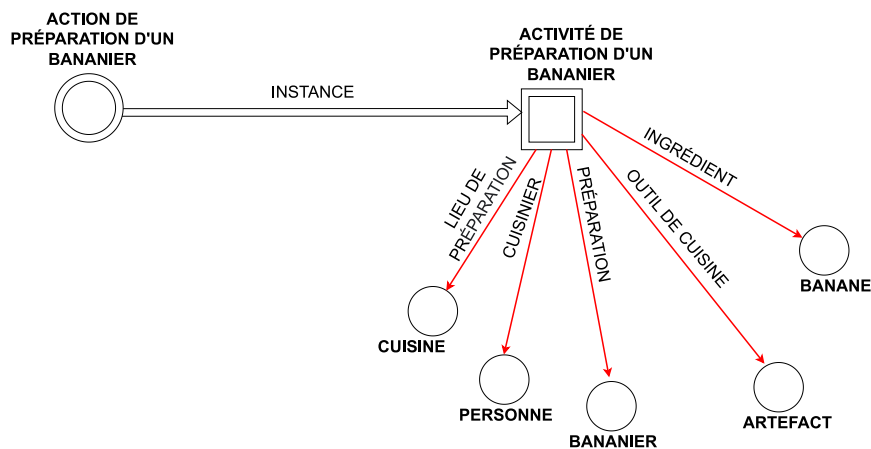


FIGURE 5.21 – FMB décrivant l'activité et l'action de préparation d'un bananier.

Pour vérifier si la propriété sémantique 'ingrédient' est la propriété sémantique pertinente pour la relation reliant BANANE à BANANIER₃, nous considérons d'autres familles basées sur des

noms d'ingrédients qui sont utilisés pour des préparations. Des couples que nous considérons comme étant pertinents sont proposés en (97). Comme on peut le voir en (97) on a d'autres couples dont l'élément initial est un N_{fru} , mais on a aussi des couples contenant des ingrédients comme ANGUILE et ANCHOIS, qui ne dénotent pas un fruit. Nous considérons que tous les couples en (97) jouent respectivement le rôle de INGRÉDIENT et de PRÉPARATION dans un scénario de préparation culinaire.

- (97) a. {FRAMBOISE.N, FRAMBOISIER.N}
 b. {FRAISE.N, FRAISIER.N}
 c. {CITRON.N, CITRONNIER.N}
 d. {POIVRON.N, POIVRONNADE.N}
 e. {ANGUILLE.N, ANGUILLADE.N}
 f. {ANCHOIS.N, ANCHOÏADE.N}

Nous considérons qu'à partir des familles qui contiennent les sous-ensembles en (97) nous pouvons également produire des histoires qui décrivent un scénario de préparation en cuisine et produire des FMBs similaires à celui en Figure 5.21. Sur cette base, nous proposons le LMB décrivant toute activité de préparation en cuisine en Figure 5.22. La projection de ce LMB sur les familles qui contiennent les sous-ensembles en (97) donne le paradigme à deux colonnes en Table 5.22. Les deux participants décrits dans ce paradigme sont tous les deux de type ontologique NOURRITURE : ce qui les distingue ce sont leurs rôles d'INGRÉDIENT et de PRÉPARATION dans le scénario de cuisine.

Comme pour d'autres LMBs que nous avons considérés précédemment, nous considérons que le LMB en Figure 5.22 peut être également obtenu à partir de familles dont l'élément initial n'est pas l'INGRÉDIENT, mais plutôt la PRÉPARATION. Des exemples de ces familles sont proposés en (98). Ces familles contiennent des noms qui jouent le rôle du CUISINER (SAUCIER, CRÊPIER), des noms d'OUTIL DE CUISINE comme CRÊPIÈRE, GAUFRIER et COUSCOUSSIER

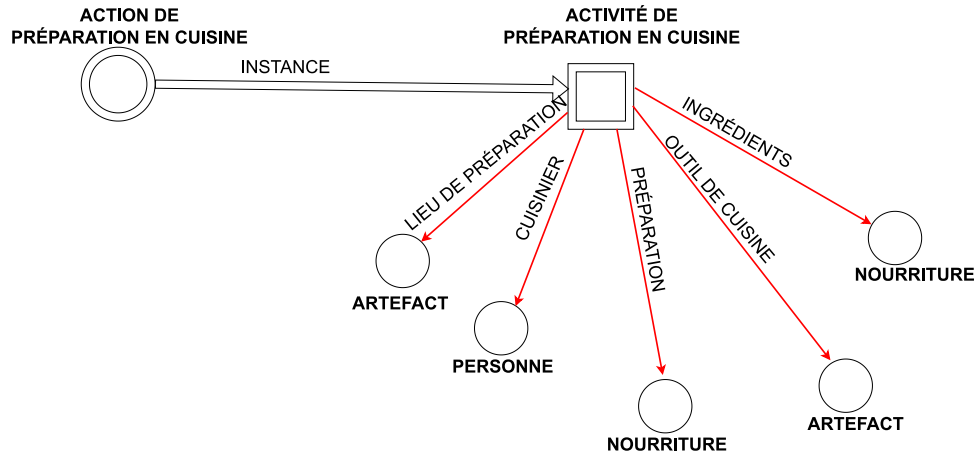


FIGURE 5.22 – LMB décrivant une activité de préparation d'un plat en cuisine.

INGRÉDIENT ENTITÉ COM.	PRÉPARATION NOURRITURE
BANANE	BANANIER
FRAMBOISE	FRAMBOISIER
ABRICOT	ABRICOTIER
CITRON	CITRONNIER
ANCHOIS	ANCHOÏADE
POIVRON	POIVRONNADE

TABLE 5.13 – Paradigme regroupant des familles qui interviennent dans une préparation culinaire.

et des noms dénotant des LIEUX DE PRÉPARATION (CRÊPERIE, COUSCOUSSERIE et CROISSANTERIE). En outre, dans la famille de CRÊPE, CRÊPERIE dénote également l'activité de faire des crêpes.

- (98) a. {CRÊPE.N, CRÊPIÈRE.N, CRÊPIER.N, CRÊPERIE.N}
 b. {COUSCOUS.N, COUSCOUSSIER.N, COUSCOUSSERIE.N}²³
 c. {GAUFRE.N, GAUFRIER.N, GAUFRERIE.N}
 d. {SAUCE.N, SAUCIER.N, SAUCIÈRE.N}²⁴

23. COUSCOUSSER.V est également attesté, mais ce verbe dénote l'action de manger du couscous plutôt que celle de le préparer.

24. SAUCIÈRE.N dénote également un récipient dans lequel on sert les sauces, mais ce sens ne rentre pas dans un scénario de préparation de sauces. SAUCERIE est attesté, mais dénote un office de saucier qui était utilisé au Moyen Âge.

e. {CROISSANT.N, CROISSANTERIE.N}

Nous considérons que la projection du LMB en Figure 5.22 sur les familles qui contiennent les sous-ensembles en (98) peut produire un deuxième paradigme dérivationnel, que nous proposons en Table 5.14. Ce paradigme montre que dans les familles dont l'élément initial est un PLAT PRÉPARÉ, on trouve régulièrement au moins un entre le CUISINIER, un USTENSILE DE CUISINE utilisé pour préparer le plat ou le LIEU DE PRÉPARATION où ce plat est cuisiné.

PLAT PRÉP. NOURRITURE	CUISINIER PERSONNE	USTENSILE DE CUISINE ARTEFACT	LIEU DE PRÉPARATION ARTEFACT	ACTIVITÉ.N
SAUCE	SAUCIER	-	-	-
CRÊPE	CRÉPIER	CRÉPIÈRE	CRÊPERIE	CRÊPERIE
GAUFRE	-	GAUFRIER	-	-
COUSCOUS	-	COUSCOUSSIER	COUSCOUSSERIE	-
CROISSANT	-	-	CROISSANTERIE	-

TABLE 5.14 – Paradigme dérivationnel de familles basées sur un nom de préparation culinaire.

Pour récapituler, à partir de la famille de BANANE, nous avons pu identifier trois familles paradigmatiques distinctes qui sont pertinentes pour trois propriétés distinctes qui font partie du sens lexical de BANANE : à savoir son rôle de PRODUIT dans le cadre d'une activité de production fruiticole, son rôle d'INGRÉDIENT dans un scénario de fabrication et son rôle de MARCHANDISE dans un scénario de transport spécialisé. En outre, nous avons vu qu'un scénario de préparation en cuisine peut être évoqué par des familles dont l'élément initial joue le rôle d'INGRÉDIENT ou de PLAT CUISINÉ.

5.6 Exemples de relation entre LMBs

Pour conclure ce chapitre, nous abordons une dernière question que nous avons anticipée dans le Chapitre 3, à savoir la possibilité de relier les réseaux sémantiques qui structurent les paradigmes via des relations “inter-réseaux”, de manière similaire à ce qui est fait dans *FrameNet*, dans lequel les cadres peuvent être reliés par des relations de nature variée.

Pour faire un exemple de relation que nous pouvons établir entre des LMBs, nous considérons que le LMB décrivant les relations entre les exemplaires d'une même espèce animale que nous

avons proposé en Section 5.2.1 et celui qui décrit une activité de chasse, qui a été présenté dans la Section 5.3.2, peuvent être connectés au moyen de relations EST-UN. Tout en contenant des participants qui ne sont pas interprédictibles, nous considérons que ce qui permet de relier ces deux LMBs est le fait qu'un animal sauvage, comme le loup, est compatible sur le plan sémantique avec les deux LMBs, car il s'agit d'une espèce animale sociale qui vit en groupe et en même temps, il s'agit d'un animal qui est chassé.

Sur cette base, les deux LMBs que nous venons de citer peuvent être reliés par une relation de type EST-UN qui connecte les trois participants du LMB décrivant les relations sociales entre exemplaires d'animaux sauvages et le participant correspondant à l'ANIMAL CHASSÉ dans un LMB qui décrit une activité de chasse. Cette relation est représentée en Figure 5.23. En outre, une autre relation de même type peut être utilisée pour relier le LMB décrivant la relation entre exemplaires de la même espèce animale sauvage et le LMB qui décrit l'activité d'étude d'un animal, qui a été discuté en Section 5.3.4.

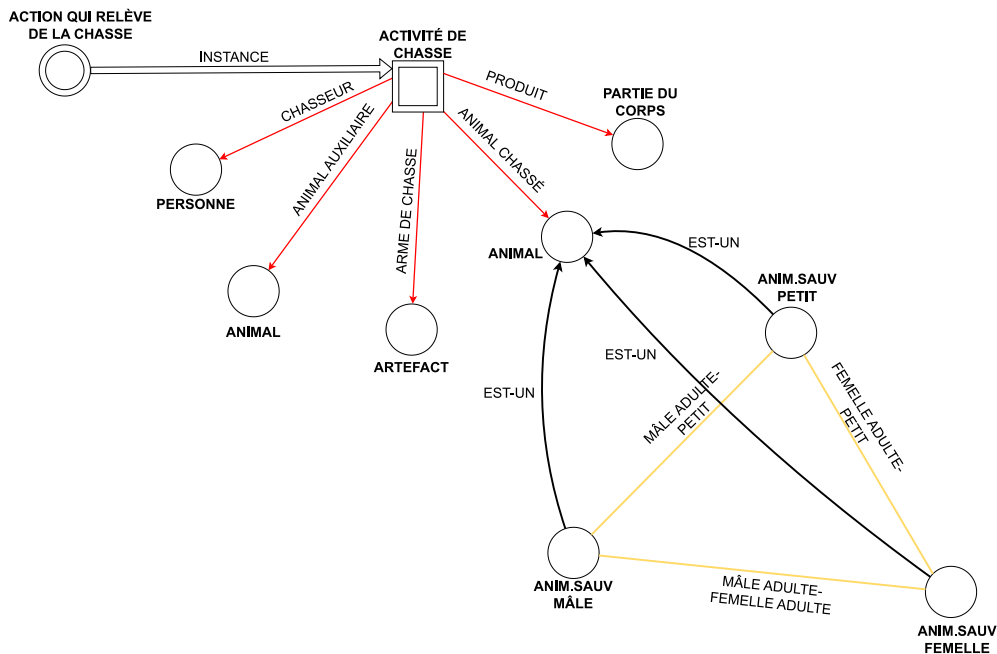


FIGURE 5.23 – Exemple de relation inter-LMBs : les exemplaires d'animal sauvage du LMB décrivant les relations sociales entre animaux sont reliés par une relation EST-UN à l'ANIMAL CHASSÉ du LMB décrivant l'activité de chasse.

Suivant le même principe, le LMB qui décrit les relations entre un insecte, son nid et sa colonie peut être relié au LMB qui décrit la relation de prédation entre un animal sauvage PROIE et un animal sauvage PRÉDATEUR. Cette relation entre deux LMBs, que nous représentons en Figure 5.24, peut être exemplifiée par le cas de la fourmi, qui sémantiquement rentre bien dans le LMB décrivant la vie sociale des insectes et rentre aussi dans le LMB décrivant la prédation en tant que PROIE par rapport à des animaux comme le fourmilier, comme nous avons pu le voir en Section 5.2.3.

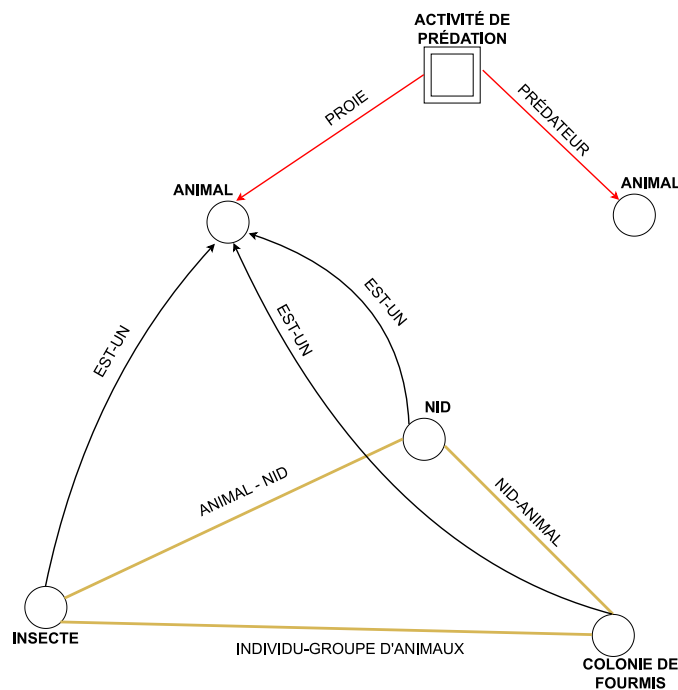


FIGURE 5.24 – Relations reliant un LMB décrivant la vie sociale des insectes et un LMB décrivant une activité de prédation de la part d'un animal

De manière générale, la question de la hiérarchisation des LMBs reste à approfondir dans le cadre de ce travail et il est possible que d'autres LMBs que nous avons présentés dans cette thèse puissent être reliés et hiérarchisés. Par exemple, le LMB décrivant une activité d'élevage en masse d'animaux présenté en Section 5.3.3 et le LMB décrivant la production fruiticole présenté en Section 5.5.1 pourraient être reliés à un LMB plus générique et hiérarchiquement supérieur qui décrit à la fois la culture animale et végétale. Les deux exemples que nous venons de fournir

montrent que certains LMBs que nous avons proposés dans ce chapitre peuvent être organisés de manière hiérarchique.

5.7 Bilan

Au début de ce chapitre, notre objectif était d’approfondir le rôle des propriétés sémantiques des lexèmes dans la structuration des familles et des paradigmes dérivationnels. Nous avons considéré principalement deux catégories ontologiques, à savoir ANIMAL et FRUIT, et nous avons pris en compte un ensemble de familles basées sur N_{anim} et N_{fru} pour déterminer dans quelle mesure ces deux catégories sont pertinentes pour la structuration des familles et des paradigmes.

Dans un premier temps, nous avons pris en compte des familles basées sur N_{anim} en excluant les dérivés qui s’inscrivent dans des activités humaines. À partir du contenu des familles basées sur N_{anim} sauvages, nous avons pu identifier : **(a)** un LMB décrivant les interactions sociales entre exemplaires d’une même espèce animale, **(b)** un LMB décrivant la relation entre un insecte, sa colonie et son nid **(c)** un LMB décrivant la relation entre une espèce animale sauvage et son espèce prédatrice. Les deux LMBs se distinguent des LMBs de type activité, car ils sont structurés par des relations qui ne sont pas de nature argumentale et qui fonctionnent comme des abstractions d’événements de nature différente. Comme nous l’avons vu, ces paradigmes décrivent des régularités sémantiques qui sont soit très spécifiques et qui concernent peu de familles (interactions sociales et insectes) soit plus générales, mais peu réalisées de manière morphologique (prédation).

En outre, en prenant en compte les familles basées sur N_{anim} , nous avons pu également identifier des paradigmes relevant de domaines d’activité variés : pêche, chasse, élimination d’animaux nuisibles, élevage et étude scientifique des espèces animales. La prise en compte ces domaines d’activité a montré que les paradigmes basés sur les domaines d’activité comportent plusieurs problèmes sur le plan de la capacité prédictive : ils comportent des lexèmes relevant de la terminologie, ils peuvent comporter des composés savants, des dérivés peu attestés et de

manière générale, ils décrivent des régularités morphologiques très limitées. Parmi le peu de régularités morphologiques observables par les paradigmes que nous avons proposés dans ce chapitre, nous avons vu une présence régulière de suffixations en *-ier* réalisant des fonctions sémantiques très diverses (ex. animal prédateur, bateau de pêche, bateau spécialisé dans le transport, gardien d'animaux, etc.)

La prise en compte des familles basées sur N_{fru} , pour sa part, a fait émerger le fait que des paradigmes dérivationnels distincts peuvent sélectionner des propriétés sémantiques distinctes qui coexistent dans le sens lexical du même lexème. Nous l'avons vu avec BANANE et sa relation avec BANANIER : nous avons pu identifier trois paradigmes distincts qui s'appuient sur trois propriétés sémantiques distinctes de BANANE ('fruit', 'marchandise', 'ingrédient') et sur trois sens distincts de BANANIER. En outre, l'exemple du paradigme de transport spécialisé a montré que les propriétés sémantiques pertinentes pour la structuration des paradigmes peuvent être des propriétés très spécifiques et qui ne sont pas les plus saillantes dans le sens lexical d'un lexème. Par ailleurs, nous avons montré que le paradigme contenant les N_{fru} , les noms des plantes et de la plantation n'est pas structuré par les propriétés ontologiques que l'on peut attribuer aux fruits, mais plutôt à une activité de production fruiticole, qui correspond à l'origine prototypique des fruits.

Dans ce chapitre, nous avons pu étendre l'inventaire des LMB que l'on peut identifier en partant par le contenu des familles dérivationnelles et nous avons approfondi la question du rôle des propriétés sémantiques dans la structuration des paradigmes. Dans le Chapitre 6 nous abordons la question des relations très régulières dans le lexique construit français, mais qui pourtant ne peuvent pas être caractérisées au moyen d'histoires : il s'agit de relations que nous définissons comme étant périphériques par rapport aux LMBs.

Chapitre 6

Relations périphériques

Sommaire

6.1	Introduction	205
6.2	Couples de noms d'humain masculin - féminin	207
6.3	Adjectifs relationnels	210
6.4	Adverbes en <i>-ment</i>	214
6.5	Noms d'action	214
6.6	Réitération d'un procès	215
6.7	Bilan	217

6.1 Introduction

Comme nous l'avons montré au Chapitre 4, la démarche que nous proposons permet de décrire les relations morphosémantiques qui existent entre les lexèmes inscrits dans les familles dérivationnelles et de découper ces familles pour identifier des familles paradigmatiques. Les familles paradigmatiques sont des sous-ensembles de familles dérivationnelles qui contiennent des lexèmes interprédictibles sur le plan sémantique et qui fonctionnent ensemble : si une famille a comme élément initial un nom qui dénote un objet fabriqué, il est probable que la même famille contienne un nom dénotant le fabricant et un nom dénotant le domaine d'activité relatif à sa production.

Pour ce faire, au début de notre démarche, nous prenons en compte tous les sous-ensembles contenus dans une famille dérivationnelle de départ, car nous ne savons pas lesquels correspondent aux familles paradigmatiques que nous cherchons, et nous construisons des histoires. Comme nous l'avons vu, les relations récurrentes dans les histoires peuvent être organisées dans des *meaning bundles*, qui peuvent être structurés par des activités humaines (c.f. Chapitre 4 et 5) ou par des réseaux non argumentaux (c.f. Chapitre 5).

Le fait de prendre en compte tous les sous-ensembles d'une famille dérivationnelle complète pour construire des histoires et pour identifier les familles paradigmatiques fait que lorsque l'on considère la famille de POT, nous devons considérer également le couple {POTIER.N, POTIÈRE.N}, que dans la même famille on a aussi le couple {POT.N, POTIER.A}, que dans la famille de FABRIQUER on a le couple {FABRIQUER.V, FABRICATION.N}, que dans cette même famille nous retrouvons le couple {FABRIQUER.V, REFABRIQUER.V} et que dans la famille de ORNITHOLOGIE nous avons le couple {ORNITHOLOGIQUE.A, ORNITHOLOGIQUEMENT.R}.

Comme nous allons le voir dans ce chapitre, les relations existant entre les lexèmes de ces couples ne peuvent émerger grâce aux histoires et ne font pas partie de la structure des réseaux de type activité ou de type non argumental comme ceux que nous avons proposés précédemment dans cette thèse.

Les relations entre masculin et féminin pour les noms d'humain, entre adjectifs relationnels et leur nom base, entre verbe et nom d'action ayant le même sens, entre lexèmes ayant un sens processif et leurs dérivés en *re-* qui dénotent une réitération du même procédé et entre adjectifs et adverbes en *-ment* nécessitent un traitement différent dans le cadre de notre démarche. Dans ce chapitre, nous rendons compte du statut de ces relations en montrant en quoi elles se distinguent des relations morphosémantiques "centrales" considérées précédemment dans cette thèse. Nous considérons que les relations que nous venons de citer ne peuvent pas être inscrites dans des LMBs, mais plutôt qu'elles sont indépendantes par rapport aux scénarios et jouant un rôle périphérique. Sur cette base, nous introduisons la notion de *relation périphérique*. Nous précisons que les exemples de relations non centrales dans les LMBs que nous fournissons dans ce chapitre ne constituent pas une liste exhaustive, car nous ne connaissons pas à ce stade de ce

travail la totalité des relations qui rentrent dans cette catégorie. Les exemples fournis dans ce chapitre ont donc l'objectif d'illustrer le traitement de quelques cas de figure parmi les relations morphologiques régulières qui ne dépendent pas de la structure des LMBs.

Définition 12

Une relation périphérique est une relation morphosémantique qui ne structure pas un LMB de type argumental décrivant un domaine d'activité ni un LMB structuré par des relations non argumentales.

Relation périphérique

6.2 Couples de noms d'humain masculin - féminin

Les réseaux activité que nous avons décrits dans les Chapitres 4 et 5 impliquent systématiquement un participant humain : par exemple, dans le LMB qui décrit une activité de fabrication le référent d'un nom d'humain comme POTIER joue le rôle du FABRICANT (c.f. Chapitre 4). En revanche, dans les familles partielles que nous avons prises en compte pour identifier les LMBs activité, nous n'avons pas pris en compte les noms féminins dénotant ces métiers. Or, dans la famille de POTIER, on a aussi le nom POTIÈRE. Dans la terminologie utilisée par Bonami et Boyé (2019), POTIER et POTIÈRE sont des *gender-iconic nouns* (GiN), à savoir des noms dont le genre correspond au genre de leur référent. Au contraire, des noms comme PERSONNE ou MAUVIETTE ne sont pas des GiNs. Deux GiNs constituent ce que Bonami et Boyé (2019) appellent un *gender iconic pair* (GiP), s'ils forment un couple de GiNs morphologiquement reliés qui se distinguent uniquement sur la base du genre. Dans les GiPs on a aussi des noms d'humain qui ne réfèrent pas à des métiers, comme ADOLESCENT-ADOLESCENTE (99b) et des noms dénotant l'exemplaire mâle et femelle d'une espèce animale, comme le couple {LOUP.N, LOUVE.N} que nous avons pris en compte dans le Chapitre 5 (99c).

- (99) a. {POTIER.N, POTIÈRE.N}
 b. {ADOLESCENT.N, ADOLESCENTE.N}
 c. {LOUP.N, LOUVE.N}

Comme nous l'avons vu en Section 5.2.1, dans le cadre des dynamiques sociales de leur espèce animale, un loup et une louve peuvent jouer des rôles différents dans les histoires : par exemple, une louve allaite un louveteau alors que le loup ne va pas le faire. En revanche, nous considérons que pour POTIER et POTIÈRE, les référents de ces deux noms jouent le même rôle dans les histoires que nous construisons en partant des membres de leur famille : un potier et une potière font tous les deux des pots. Un exemple d'histoire où POTIER et POTIÈRE sont interchangeables est fourni en (100).

- (100) *Dans un petit village, [un talentueux potier / une talentueuse potière] sculptait un pot extraordinaire. Son travail était si exceptionnel que le pot devint une légende locale, célèbre bien au-delà du village. [Le potier était fier / La potière était fière] de sa création qui témoignait de son amour pour l'art de la poterie.*

Cette considération s'applique à tous les noms dénotant des métiers : un vacher et une vachère élèvent s'occupent des vaches, un pruniculteur et une prunicultrice cultivent des prunes, un dentiste et une dentiste soignent les dents, etc. En outre, elle s'applique aussi à des noms qui réfèrent à des personnes qui se consacrent à des activités qui ne sont pas forcément professionnelles : un skieur et une skieuse skient, un fumeur et une fumeuse fument, etc.

Cela signifie que les membres de GiPs qui dénotent des personnes qui se consacrent à des activités jouent le même rôle par rapport à celles-ci : le potier et la potière jouent le rôle de FABRICANT dans un LMB qui décrit une activité de fabrication et le pruniculteur et prunicultrice jouent le rôle de l'AGRICULTEUR dans un LMB de production fruiticole.

Comment décrire les relations qui existent entre les membres de ces couples ? Comme nous l'avons anticipé, nous considérons que ces relations ne peuvent pas être caractérisées au moyen d'histoires, car la distinction entre ces noms se fait uniquement sur la base du genre. Plusieurs études ont traité de la question de la nature de la relation entre les GiPs et leur statut est débattu : une première tradition considère que les GiPs sont reliés par des relations de nature dérivationnelle et que donc chaque GiP contient des lexèmes distincts (Zwanenburg, 1988; Matthews, 1974). D'autres, comme Bonami et Boyé (2019) considèrent que la nature flexionnelle

ou dérivationnelle des relations reliant les GiPs dépend de la langue. Pour les langues dans lesquelles les noms d'humain de l'un des genres sont régulièrement dérivés à partir de l'autre genre, il est raisonnable de considérer que les GiPs contiennent des lexèmes reliés par des relations dérivationnelles. En revanche, ils considèrent que cela n'est pas le cas du français. Ils observent deux groupes principaux : **(a)** GiPs contenant des noms de genre non marqué, comme DENTISTE (masculin) - DENTISTE (féminin) **(b)** GiP formés par suffixations “parallèles” et non l'un à partir de l'autre, comme POTIER et POTIÈRE, JOUEUR et JOUEUSE, AUDITEUR et AUDITRICE. La proposition de Bonami et Boyé (2019) est de considérer les relations reliant les GiPs en français comme étant de nature flexionnelle.

Pour intégrer la description de relation masculin-féminin chez les noms d'humain dont la nature dérivationnelle est remise en cause, nous considérons que les participants humains qui sont impliqués dans un LMB de type activité sont reliés à leur tour par une relation EST-UN aux deux genres possibles qui peuvent les caractériser. La relation EST-UN que nous venons d'introduire est périphérique par rapport aux LMBs : elle ne fait pas partie de la structure du scénario décrivant l'activité, mais elle est uniquement associée aux participants humains qui y sont impliqués. En outre, ces relations sont “automatiques” : à partir du moment où un LMB contient un participant humain, ce participant humain peut forcément se réaliser au masculin ou au féminin dans le lexique.

Pour résumer, les référents des membres d'un GiP qui dénotent des humains correspondent au même participant dans un LMB et sont distingués uniquement au niveau du genre. La présence de deux réalisations possibles qui se distinguent uniquement au niveau du genre est systématique pour chaque participant humain qui est impliqué dans un LMB. Pour les animaux, en revanche, nous avons vu que le mâle et la femelle peuvent être distingués aussi par d'autres facteurs que le genre et avoir un rôle distinct dans les histoires. Lorsque nous projetons un LMB sur une famille dérivationnelle pour la découper, les lexèmes faisant partie d'un GiP qui comprend des noms d'humain font partie de la même famille paradigmatique et remplissent deux colonnes distinctes dans le paradigme. Ces colonnes se distinguent uniquement sur le plan du genre, comme il est montré en Table 6.1.

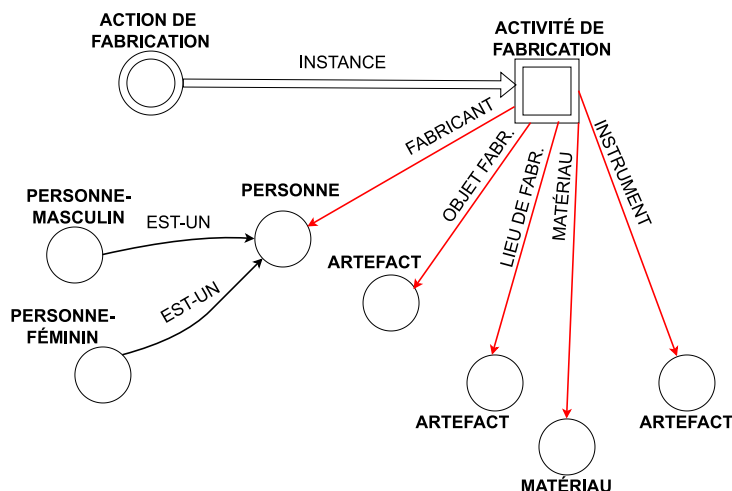


FIGURE 6.1 – Chaque participant humain impliqué dans un LMB est relié par des relations périphériques de type EST-UN aux deux genres possibles qui le caractérisent. Nous attribuons les étiquettes PERSONNE-MASCULIN et PERSONNE-FÉMININ au deux sommets qui représentent les deux genres. La présence de ces relations est systématique pour chaque participant humain impliqué dans un LMB.

OBJET FABRIQUÉ ARTEFACT	FABRICANT PERSONNE.M	FABRICANT PERSONNE.F	LIEU DE FABRICATION ARTEFACT	ACTIVITÉ.N
POT	POTIER	POTIÈRE	POTERIE	POTERIE
BRIQUE	BRIQUETIER	BRIQUETIÈRE	BRIQUETERIE	BRIQUETERIE
...

TABLE 6.1 – Dans un paradigme dérivationnel, les membres d’un GiP sont réalisés dans deux colonnes distinctes qui se distinguent uniquement au niveau du genre (dans ce cas, la deuxième et la troisième colonne de la table).

6.3 Adjectifs relationnels

Plusieurs familles considérées dans les Chapitres 4 et 5 comportent des adjectifs relationnels (Adj_{rel}). Ces adjectifs ont fait l’objet de plusieurs travaux (McNally et Boleda, 2004; Fradin, 2007; Strnadová, 2014; Namer, 2021). Les adjectifs relationnels sont des adjectifs dénominatifs qui partagent leur contenu sémantique avec leur nom de base. Par exemple, POTIER est un adjectif en (101). Les Adj_{rels} établissent une relation entre leur nom de base et le nom qu’ils modifient dans un syntagme nominal (Fradin, 2007; Namer, 2021). Par exemple, en (101) POTIÈRE est associée à *production*. Les adjectifs relationnels sont sémantiquement équivalents à un syntagme prépositionnel ayant une structure $De+N_{base}$: *la production potière* équivaut donc à *la production*

des pots.

(101) *Lezoux, dans le département du Puy-de-Dôme, a été l'un des plus grands centres de production potière de l'Empire romain.*

Étant donné que le contenu sémantique des Adj_{rels} est équivalent à celui de leur nom de base, nous considérons que la relation entre un Adj_{rel} et sa base nominale ne peut pas être caractérisée au moyen d'histoires et ne fait pas partie des relations qui structurent les LMBs. Il s'agit plutôt d'un autre type de relation périphérique par rapport aux LMBs : les adjectifs relationnels sont des éléments annexes qui s'associent aux participants des LMBs et qui y sont reliés par une relation de TRANSPOSITION. En outre, nous avons vu dans le Chapitre 2 que plusieurs études ont montré que les adjectifs relationnels peuvent souvent avoir plusieurs interprétations. C'est le cas par exemple de JOURNALISTIQUE, qui est construit sur JOURNALISTE, mais qui peut être également associé à JOURNAL et à JOURNALISME, ou encore de ARTISANAL qui peut s'associer à ARTISAN ou à ARTISANAT (Strnadová, 2014; Namer, 2021). Des exemples des trois emplois de JOURNALISTIQUE sont proposés en (102) (Namer, 2021, p.122).

- (102) a. *Rédigé en vers, le texte à la typographie journalistique est illustré de dessins et de collages.* = typographie des journaux
 b. *le Souverain exprime aux membres de la famille du défunt et, à travers eux, à sa grande famille journalistique et médiatique.* = famille des journalistes
 c. *l'introduction de la presse électronique a estompé les frontières entre les différents métiers journalistiques, de la rédaction à la fabrication.* = métiers du journalisme

Les Adj_{rels} à interprétation multiple peuvent être associés à plusieurs lexèmes faisant partie de la même famille. Pour expliquer ces interprétations multiples, plusieurs auteurs parlent d'un principe d'économie : selon ce principe, la langue, plutôt que de créer une nouvelle forme, utilise une forme déjà existante (Roché, 2008, 2009; Strnadová, 2014).

Cela signifie que les sens possibles que l'on peut associer à un adjectif relationnel dépendent **(a)** du contenu de sa famille, **(b)** du principe d'économie que nous venons de citer. Namer

(2021) montre que l'interprétation multiple des adjectifs relationnels est fortement corrélée à la nature sémantique des relations qui impliquent le nom base et que, lorsque ces relations sont centrées sur l'humain et sur les activités humaines, les Adj_{rels} ont souvent des interprétations multiples. ÉLECTORAL est associable à ÉLECTEUR et à ÉLECTION, SÉNATORIAL est associable à SÉNAT et à SÉNATEUR, ATHLÉTIQUE est associable à ATHLÈTE et à ATHLÉTISME, etc. Namer (2021) a montré que les adjectifs relationnels à interprétations multiple peuvent s'associer, entre beaucoup d'autres, à des noms d'agents qui se consacrent à une activité (103a), à des activités (103b), à des domaines de connaissance (103c), à des instruments qui sont utilisés dans des activités (103d) et à des noms de doctrines (103e).

- (103) a. *génération accordéonistique = génération d'accordéonistes*
 b. *processus électoral = processus de l'élection*
 c. *principe spectrographique = principe de la spectrographie*
 d. *tracé spectrographique = tracé du spectrographe*
 e. *fondement patriotique = fondement du patriotisme*

Sur cette base, nous considérons que dans les LMB de type activité, l'activité et les participants qui potentiellement peuvent être associés à un Adj_{rel} sont reliés par une relation périphérique à un adjectif relationnel. Nous représentons ces relations par des arêtes pointillées, comme en Figure 6.2, qui symbolisent une relation de type TRANSPOSITION. Comme on peut le voir dans cette Figure, nous considérons que la relation de transposition est presque systématique. Comme il est montré dans les exemples en (104), pour presque tous les participants du LMB en Figure 6.2 on peut trouver un lexème qui les réalise associable à un Adj_{rel} . Dans le cadre de l'activité de fabrication schématisée en Figure 6.2, le seul participant auquel nous n'associons pas est l'INSTRUMENT DE FABRICATION : bien que les Adj_{rel} peuvent s'associer à des instruments (ex. *lentille otoscopique*) il semble plus difficile de trouver un Adj_{rel} associé à un instrument de fabrication (ex. TOUR, MARTEAU, SOUDEUSE, TOURNEVIS) mais, sur le plan sémantique, rien empêche d'avoir un Adj_{rel} associé à ce sommet du LMB.

- (104) a. l'Adj_{rel} POTIER peut être associé à POTERIE (ex. *activité potière*) ou à POT (ex. *fabrication potière*);
- b. l'Adj_{rel} ARTISANAL peut être associé à ARTISANAT (ex. *activité artisanale*) ou à ARTISAN (*syndicat artisanal des métiers de l'art et de la création*);
- c. l'Adj_{rel} USINIER est associé à USINE (ex. *la vie usinière de la région est en crise*), qui est un LIEU DE FABRICATION;
- d. l'Adj_{rel} COTONNIER est associé à COTON (ex. *tissu cotonnier*), qui peut intervenir en tant que MATÉRIAU dans un scénario de fabrication.

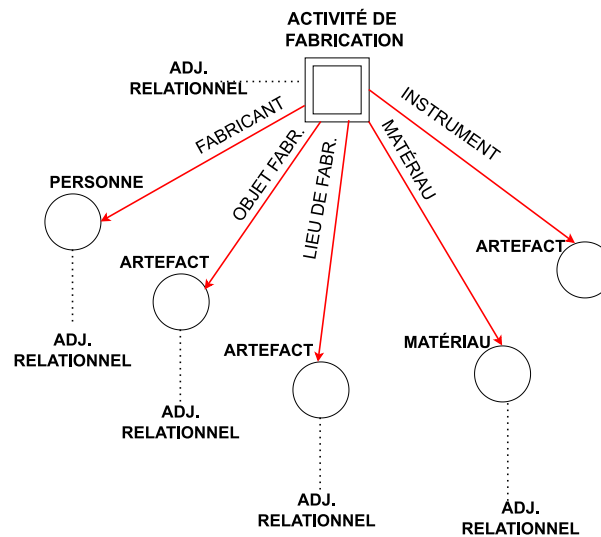


FIGURE 6.2 – La relation entre participants d'un LMB et leur Adj_{rel} est représentée au moyen d'une arête pointillée. Tous les participants du LMB qui peuvent être associés à un Adj_{rel} sont reliés à un Adj_{rel} dans le schéma. Le seul participant pour lequel nous n'avons pas d'exemple de Adj_{rel} associable est l'INSTRUMENT DE FABRICATION.

Pour résumer, la relation entre les Adj_{rels} et leur nom de base est un autre type de relation périphérique par rapport aux LMB. Il s'agit d'une relation de type TRANSPOSITION existant entre lexèmes dénotant le même sens et qui se distinguent uniquement sur le plan catégoriel.

6.4 Adverbes en *-ment*

Dans cette section, nous abordons la question des adverbes en *-ment*. Les adverbes en *-ment* peuvent se former à partir d’adjectifs qualificatifs qui expriment une propriété (ex. SALE.A - SALEMENT.R) ou à partir d’adjectifs relationnels. Par exemple, dans la famille de ENTOMOLOGIE l’Adj_{rel} ENTOMOLOGIQUE et l’adverbe ENTOMOLOGIQUEMENT. Le statut de lexème des adverbes en *-ment* a été questionné par Dal (2007, 2018), qui propose de les considérer comme des mots formes d’adjectifs. Selon cette proposition, les adverbes en *-ment* relèveraient de la flexion adjectivale et non de la dérivation.

Nous adoptons la proposition de Dal (2007) et, pour intégrer les adverbes en *-ment* à notre méthode, nous les considérons comme étant des éléments annexes aux adjectifs par une relation périphérique de type flexionnel. Dans le cas des adverbes construits sur des Adj_{rels}, ils sont des éléments annexes aux Adj_{rels}, qui à leur tour sont des annexes des participants du LMB, comme nous l’avons vu en Section 6.3. Cela est représenté en Figure 6.3 avec le LMB qui représente une activité d’étude d’un animal que nous avons proposé en Section 5.3.4. Nous représentons les relations entre adjectifs et adverbes en *-ment* en utilisant une double ligne pour les distinguer des relations de transposition introduites dans la section précédente. Dans le LMB décrivant une activité d’étude d’un animal, on a (au moins) une relation entre un adverbe en *-ment* et l’Adj_{rel} annexe à l’activité : par exemple ORNITHOLOGIQUEMENT est annexe à ORNITHOLOGIQUE, qui à son tour signifie ‘relatif à l’ornithologie’.

6.5 Noms d’action

Pour décrire les relations entre les verbes ayant un sens processif et des noms d’action exprimant le même sens, comme dans le couple {FABRIQUER.V, FABRICATION.N}, nous proposons un traitement similaire à celui que nous avons présenté pour les adjectifs relationnels en Section 6.3. De la même manière que la relation entre Adj_{rel} et nom, les verbes et les noms dénotant le

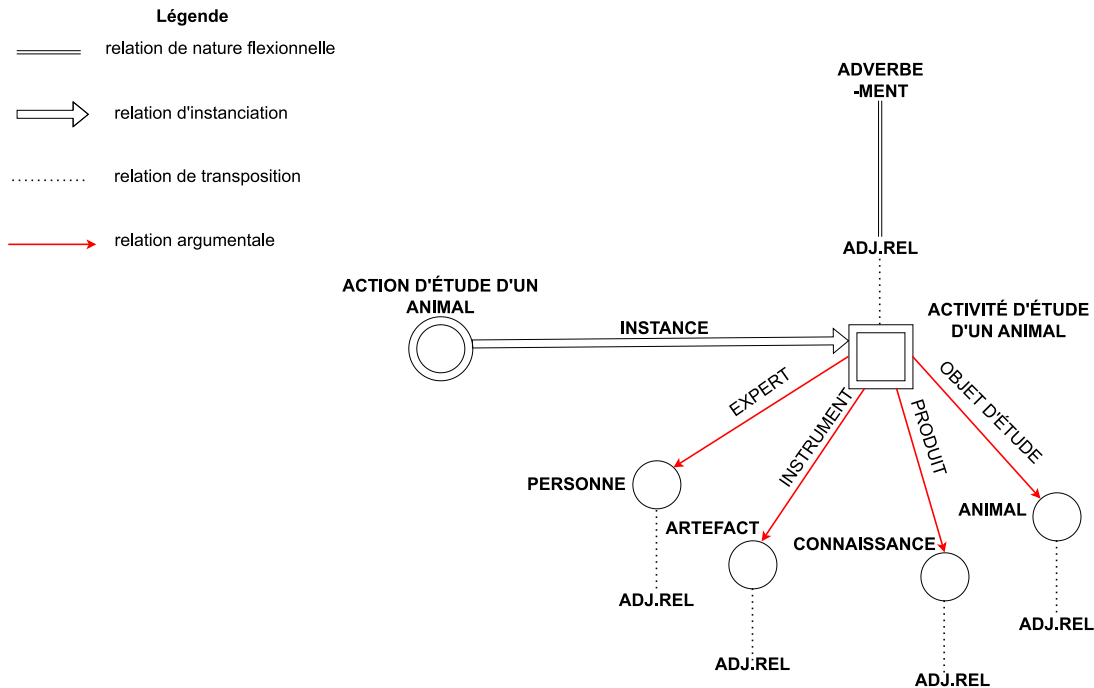


FIGURE 6.3 – Exemple d'intégration à un LMB d'un adverbe en *-ment* construit sur un Adj_{rel} par une relation périphérique de nature flexionnelle.

même procès sont reliés par une relation de type TRANSPOSITION. Cette relation est représentée par une arête pointillée et relie le pivot sémantique d'un réseau action à sa réalisation sous forme nominale, comme il est montré en Figure 6.4.

6.6 Réitération d'un procès

Dans les familles qui contiennent des verbes et des noms ayant un sens processif, il est fréquent de trouver des verbes et des noms construits qui dénotent une réitération du procédé dénoté par le nom ou verbe de base. Cette fonction sémantique en français est généralement réalisée au moyen de préfixations en *re-* (Amiot, 2002; Apothéloz, 2007). Par exemple, dans la même famille qui contient *LAVÉ* et *LAVAGE*, on a également *RELAVÉ* et *RELAVAGE*. Cela signifie que dans la famille de *LAVÉ* on a des sous-ensembles comme $\{\text{LAVÉ.V}, \text{RELAVÉ.V}\}$ ou $\{\text{LAVAGE.N}, \text{RELAVAGE.N}\}$ dont la description doit être intégrée à notre démarche.

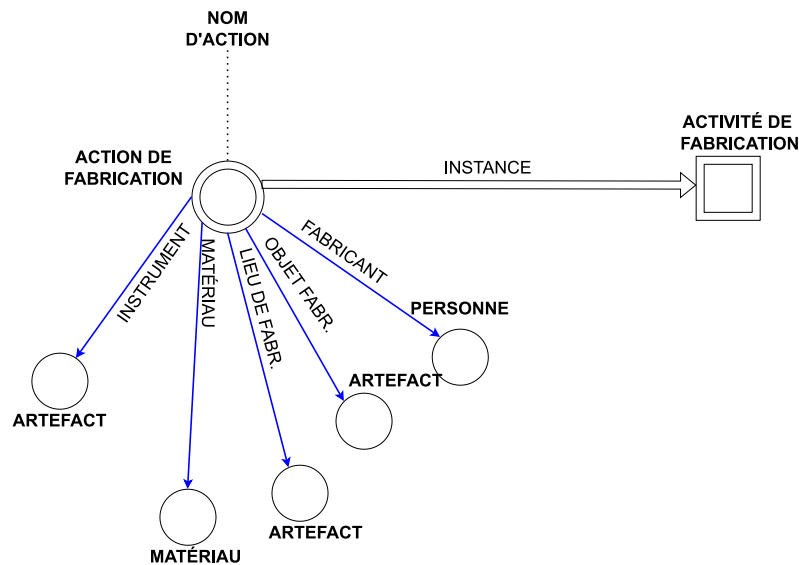


FIGURE 6.4 – Dans un réseau action, le nom d’action est relié à son équivalent verbal au sommet qui représente l’action par une relation de transposition.

Lorsqu’un lexème exprime la réitération d’un procès, on présuppose qu’une première occurrence du procès désigné par le verbe ou le nom de base a eu lieu et s’est terminée (Amiot, 2002). Le sens de réitération peut donc s’appliquer à des actions de type télique et peut se situer uniquement sur le plan action, qui décrit des procédés délimités dans le temps et dans l’espace. Sur cette base, nous considérons qu’un verbe comme RELAVER introduit un événement de même type que celui qui est introduit par le verbe LAVER et qui comporte les mêmes types de participants.

Nous considérons que la relation entre un lexème ayant un sens processif et un dérivé qui exprime une réitération d’un même procès représente un autre type de relation périphérique. Pour cette relation, nous utilisons le type relationnel RÉITÉRATION. Nous représentons cette relation au moyen d’un arc orienté. Cette relation connecte la composante action d’un LMB avec un autre réseau action qui décrit le même type d’événement, comme on peut le voir en Figure 6.5. Comme nous l’avons montré en Section 6.5, l’action peut se réaliser sous forme nominale et une action sous forme nominale est aussi susceptible de produire un dérivé avec sens itératif (ex. REFABRICATION, RECONSTRUCTION, etc.).

Comme le montre la Figure 6.5, nous considérons qu’une seule réitération d’une action ne

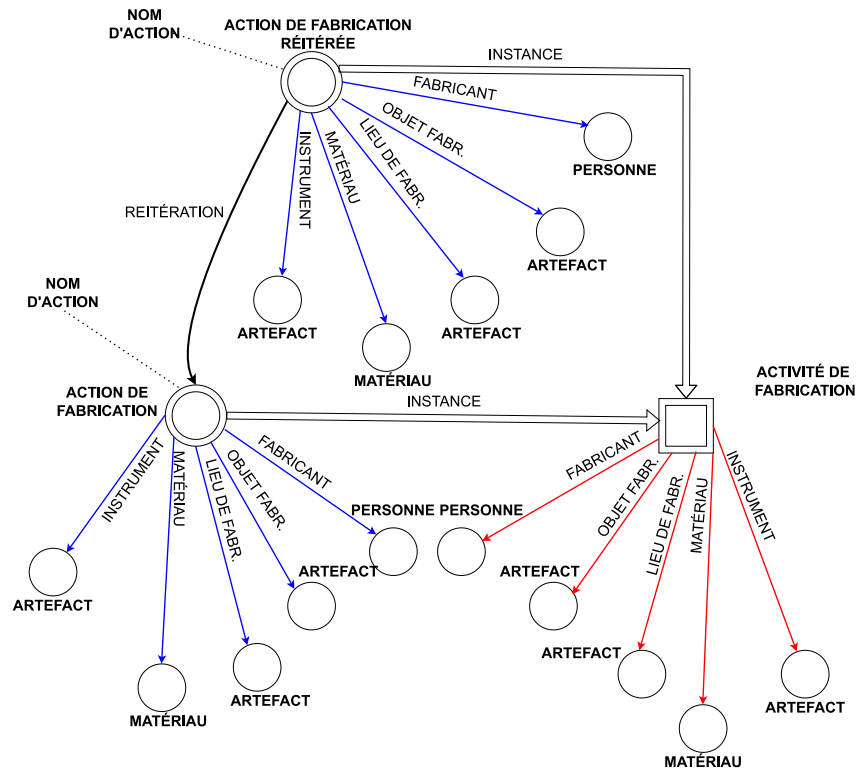


FIGURE 6.5 – Nous considérons les relations entre un prédicat et un dérivé qui dénote une réitération du même prédicat comme étant des relations périphériques par rapport aux LMB. Ces relations sont représentées par un arc orienté typé ITÉRATION. L'action et sa réitération sont toutes les deux reliés à une activité par une relation d'instanciation, car elles sont toutes les deux des instances de l'activité correspondante.

correspond pas à une activité : si on doit décrire l'activité d'un artisan, on ne dira pas que cet artisan refabrique des pots, mais plutôt qu'il fabrique régulièrement des pots. En revanche, les deux composantes action (action et sa réitération) sont toutes les deux des instances d'une activité régulière, comme le montrent les deux relations de type INSTANCE en Figure 6.5.

6.7 Bilan

Dans ce chapitre nous avons abordé la question du traitement de relations morphosémantiques que nous considérons comme étant périphériques par rapport aux LMBs. Contrairement aux relations que nous avons traitées dans les Chapitres 4 et 5, ces relations ne font pas partie

de la structure des LMBs et ont un statut secondaire par rapport aux scénarios qui sont décrits dans les LMBs. Nous avons identifié plusieurs types de relations périphériques divers : en premier lieu, nous avons abordé la description de la relation entre noms d'humain au masculin et au féminin, en nous concentrant sur des exemples contenant des noms de métiers. Dans le cadre d'une activité, nous considérons que les référents des noms au masculin et au féminin jouent le même rôle et que la relation qui les relie se situe entre flexion et dérivation. Au contraire, au Chapitre 5 nous avons vu que pour les animaux, le mâle et la femelle peuvent jouer des rôles distincts dans les histoires. La relation masculin-féminin est périphérique mais systématique dans les LMBs : chaque participant humain impliqué dans un LMB est relié par une relation de type EST-UN aux deux genres sous lesquels il peut se "manifester" dans le lexique.

Nous avons ensuite considéré la relation entre adjectifs de relation et les noms sur lesquels ils sont construits. Comme le sens d'un Adj_{rel} est le même que celui du nom auquel il est associé, nous considérons que nom et Adj_{rel} sont reliés par une relation de transposition qui ne fait pas partie de la structure du LMB. En outre, nous avons vu que lorsque le nom de base s'inscrit dans le cadre d'une activité humaine, les Adj_{rel} peuvent avoir des interprétations multiples et peuvent référer à plusieurs membres de la famille. Sur cette base, plusieurs participants d'un LMB sont susceptibles d'être associés à un Adj_{rel} . Nous considérons que verbes et noms d'action qui dénotent le même procès sont également reliés par une relation de transposition. En outre, tout en n'ayant pas traité les noms de propriété dans cette thèse, nous considérons que le type TRANSPOSITION pourrait être employé également pour décrire les relations existant entre adjectifs qualificatifs (ex. BEAU) et noms de propriété dénotant le même sens (ex. BEAUTÉ).

Ensuite, nous avons pris en compte les adverbes dérivés en *-ment* à partir d'un adjectif. En fonction de la nature quasi flexionnelle de ces dérivés, nous considérons que la relation entre adverbes en *-ment* et adjectifs relationnels ou qualificatifs fait aussi partie des relations périphériques par rapport aux LMB. En suivant Dal (2007), nous considérons qu'il s'agit d'une relation flexionnelle.

En dernier lieu, nous avons pris en compte la relation entre noms et verbes ayant un sens processif et leurs dérivés obtenus avec préfixation en *re-* qui expriment une répétition du même

procédé. Il s'agit d'une autre relation périphérique par rapport aux LMBs : elle s'établit entre le centre sémantique d'un réseau action et le centre sémantique d'un autre réseau action ayant la même structure.

Nous précisons que la gamme de relations périphériques que nous avons présentées dans ce chapitre ne couvre pas la totalité des relations hors LMBs qui peuvent être identifiées dans les familles. Par exemples, les adjectifs dérivés en *-able* sont une autre catégorie de lexèmes construits dont le statut par rapport aux LMB devrait être clarifié. Il en va de même pour les préfixations en *pré-*, qui peuvent être réalisées sous forme verbale (ex. PRÉLAVÉ), nominale (ex. PRÉLAVAGE) ou adjectivale (ex. PRÉCUIT). Pour conclure, il faudrait également rendre compte de la présence de diminutifs et évaluatifs, car par exemple dans la famille de BOTTE on trouve également BOTTINE et dans la famille de COUTEAU on trouve également COULETEL. Par exemple, il est possible que dans un scénario de fabrication de l'objet le référent d'un nom comme COULETEL joue le même rôle du référent du nom COUTEAU (tous les deux sont reliés par la même relation à COUTELIER), mais cela est moins envisageable pour un scénario d'usage, car les deux objets n'ont pas tout à fait les mêmes fonctions. Ces questions restent à approfondir.

Conclusions et perspectives

Bilan général

Dans cette thèse, nous avons adopté une approche paradigmatique à la description du lexique construit français. Notre objectif principal était de proposer une méthodologie théorique capable d’apporter une réponse au problème de la délimitation des familles et des paradigmes en dérivation. Deux hypothèses principales ont guidé ce travail : **(a)** la morphologie dérivationnelle est paradigmatique ; **(b)** les relations sémantiques entre lexèmes sont premières pour la structuration des paradigmes dérivationnels.

La méthode que nous proposons considère que les familles sont structurées par des réseaux conceptuels sous-jacents. Pour délimiter ces réseaux conceptuels, nous nous sommes inspirés de la sémantique des cadres de Fillmore (Fillmore *et al.*, 1976; Petruck, 1996), qui se sert de scénarios cognitifs pour l’annotation sémantique de corpus. Au sein de ces scénarios conceptuels, les participants sont interprédictibles et fonctionnent de manière intégrée. La notion de *meaning bundle* (MB) que nous avons proposée se base également sur ce principe d’interprédictibilité des participants au sein d’un même scénario. Nous avons proposé d’obtenir les MBs à partir d’histoires. Les histoires sont un outil heuristique que nous avons utilisé pour expliciter les relations existant entre les membres d’une famille dérivationnelle. Les MBs ne retiennent que le contenu régulier des histoires et filtrent tout ce qui est épisodique.

La méthode que nous proposons permet de découper une famille dérivationnelle complète en familles paradigmatiques. Nous avons vu que les familles paradigmatiques peuvent se chevaucher lorsqu’elles partagent un lexème. Par exemple, nous avons produit deux familles paradigmatiques produites à partir de la famille de POT : une s’inscrit dans un scénario d’origine du pot ({

POT.N, POTIER.N, POTERIE.N}) et l'autre s'inscrit dans un scénario d'usage d'un pot ({POT.N, REMPOTER.N, REMPOTEUR.N}). Quand des familles paradigmatiques partagent un lexème, elles s'appuient sur des propriétés sémantiques de celui-ci : par exemple, dans le cas de POT ces propriétés sont sa nature d'objet fabriqué et sa fonction de récipient (voir Chapitre 4).

Ce multifacettage sémantique des unités du lexique nous permet d'ajuster le niveau de structuration des paradigmes, ce qui nous a amené à montrer que les propriétés pertinentes ne sont pas toujours les plus saillantes que l'on peut associer à un lexème donné : par exemple la famille paradigmatique {BANANE.N, BANANIER.N (cargo)} s'inscrit dans un paradigme de transport spécialisé qui comprend également des dérivés comme PETROLIER (cargo qui transporte du pétrole), GAZIER (cargo qui transporte du gaz) et ASPHALTIER (cargo qui transporte de l'asphalte). Pour ce paradigme, la propriété sémantique de BANANE qui est pertinente n'est pas sa nature de fruit cultivé ni son usage dans des recettes, mais plutôt le fait qu'il s'agit d'une marchandise qui a besoin d'un transport spécialisé et de navires équipés *ad-hoc* pour son transport (voir Chapitre 5).

Nous avons également montré que la projection d'un *meaning bundle* sur les familles peut produire plusieurs paradigmes. Nous avons vu cela avec le scénario de fabrication au Chapitre 4 qui peut produire trois paradigmes distincts : un qui contient des familles paradigmatiques basées sur un nom d'un objet fabriqué (ex. {BOTTE.N, BOTTIER.N, BOTTERIE.N}), un deuxième qui contient des familles basées sur un nom de matériau utilisé dans la fabrication (ex. {BRONZE.N, BRONZIER.N, BRONZIERIE.N}) et un troisième contenant des familles basées sur un nom dénotant un instrument utilisé dans des procédés de fabrication ({TOUR.N, TOURNEUR.N, TOURNAGE.N}). Nous avons pu voir cela également au Chapitre 5 avec les scénarios d'activité de pêche et de préparation en cuisine. Cela montre que deux familles paradigmatiques qui peuvent être inscrites dans le même *meaning bundle* ne font pas nécessairement partie du même paradigme.

Dans le Chapitre 5 nous avons utilisé des scénarios correspondant à des domaines d'activité pour former des paradigmes dérivationnels. Nous avons vu que souvent les paradigmes basés sur les domaines d'activité contiennent des lexèmes appartenant à des lexiques spécialisés et qui

sont peu accessibles à un locuteur qui n'a pas de connaissance dans le domaine en question et qui sont peu attestés. Le choix de considérer que les paradigmes sont structurés par des domaines d'activité comporte donc d'avoir des paradigmes remplis par des lexèmes à très basse fréquence en termes d'attestation. En outre, nous avons vu que certains domaines spécialisés comme l'élevage et la production fruiticole emploient de manière régulière des composés de type *Xculteur* et *Xculture* pour dénoter des producteurs de fruits et des éleveurs (*Xculteur*) et des activités de cultivation ou d'élevage (*Xculture*). Nous avons proposé de considérer les composantes *-culteur* et *-culture* comme étant proches des affixes et de les intégrer aux paradigmes dérivationnels. Il en va de même pour la composante *-cide* dans les composés qui dénotent un produit pour éliminer les animaux nuisibles et pour les composantes *-logue* et *-logie* qui créent des noms de spécialiste et des noms de domaine.

En outre, nous avons vu que la plupart des paradigmes que nous avons proposés sont très spécifiques en termes de scénario sous-jacent. Nous avons fait ce choix pour assurer un alignement sémantique assez stable au niveau des familles paradigmatiques. Cela comporte que dans certains cas, on obtient des paradigmes très réduits en termes de familles paradigmatiques qui y rentrent. Un cas plutôt représentatif est celui des animaux et du paradigme qui contient les insectes, leur nid et leur colonie (ex. {FOURMI.N, FOURMILIER.N}, {GUÊPE.N, GUÊPIER.N}, {TERMITE.N, TERMITIÈRE.N}). Dans d'autres cas encore, les régularités sémantiques sont moins spécifiques mais elles sont réalisées par très peu de familles, comme dans le cas des couples PROIE-PRÉDATEUR chez les noms d'animal.

En dernier lieu, nous avons montré que les paradigmes dérivationnels ne contiennent pas seulement des dérivés que l'on peut inscrire dans le même scénario, mais aussi des dérivés qui sont impliqués dans des relations que nous avons définies comme étant périphériques par rapport aux scénarios conceptuels et qui existent indépendamment des *meaning bundles*. Il s'agit de relations comme la relation masculin-féminin chez les noms d'humain, de la relation entre adjectifs relationnels et le nom auquel ils sont associés, de la relation entre verbes et noms d'action, de la relation lexèmes ayant un sens prédicatif et leur dérivé en *re-* et de la relation entre adverbes en *-ment* et l'adjectif sur lequel ils sont construits.

Limites et perspectives de ce travail

Comme nous l'avons annoncé au cours de la thèse, la méthode que nous avons proposée est une méthode théorique. Les histoires que nous avons proposées dans cette thèse ont servi à illustrer son fonctionnement et ses différentes étapes. Une première chose à faire serait d'implémenter cette méthode en récoltant une grande quantité d'histoires construites à partir de familles. Cela nous permettrait de constituer les MBs dont nous avons besoin pour délimiter les paradigmes dérivationnels. Nous avons montré dans le Chapitre 4 que les histoires générées par des modèles génératifs comme *GPT* sont pertinentes pour nos besoins : elles sont sémantiquement cohérentes et impliquent des participants qui sont interprétables par rapport aux référents des lexèmes que l'on indique dans le prompt. Tout en reconnaissant les biais qui caractérisent les modèles génératifs de langage, nous considérons que leur usage permettrait d'avoir de manière relativement simple une grande quantité d'histoires que, une fois validées par des humains, nous permettrait de constituer les MBs que nous cherchons à produire.

Une solution alternative serait de faire produire les histoires dont nous avons besoin par des locuteurs humains. En considérant le grand nombre d'histoires dont nous avons besoin et de relations à décrire, cette solution serait sûrement moins rapide que la précédente. Si on choisissait cette piste, il serait probablement nécessaire de mettre en place une plate-forme via laquelle récolter les histoires proposées par les locuteurs, de manière similaire à ce qui a été fait avec *JeuxDeMots* (Lafourcade et Joubert, 2008).

En outre, dans cette thèse, nous avons pris en compte un petit échantillon de familles pour décrire le fonctionnement de notre méthode. Une étude plus à grande échelle devrait être faite sur le lexique construit français, de manière à valider de manière plus extensive la méthode que nous avons proposée. Pour avoir un inventaire de familles dérivationnelles de départ plus à grande échelle on pourrait s'appuyer sur *Démonette* (Hathout et Namer, 2014) et voir quelles familles paradigmatiques et paradigmes peuvent être identifiés à partir des familles contenues dans cette base de données. En outre, en restant sur *Démonette*, la méthode que nous avons proposée dans cette thèse pourrait être exploitée pour l'annotation sémantique des relations

entre les couples de dérivés qui sont présents dans cette base de données.

En dernier lieu, il serait intéressant d'appliquer la méthode proposée dans cette thèse sur des familles dérivationnelles appartenant au lexique d'autres langues, de manière similaire à ce qui a été fait par Körtvélyessy *et al.* (2020). Cela permettrait de voir dans quelle mesure les *meaning bundles* que l'on produit à partir des familles varient ou s'ils restent relativement stables et de voir le degré de réalisation morphologique des *meaning bundles* d'une langue à une autre.

Bibliographie

- ACKERMAN, F., BLEVINS, J. P. et MALOUF, R. (2009). Parts and wholes : Implicative patterns in inflectional paradigms. *Analogy in grammar : Form and acquisition*, 54:81.
- AMIOT, D. (2002). Re-, préfixe aspectuel? *In Temps et aspect*, pages 1–20. Brill.
- AMIOT, D. et DAL, G. (2008). La composition néoclassique en français et l'ordre des constituants. *La composition dans une perspective typologique. Arras : Artois Presses Université*, pages 89–113.
- ANDERSON, S. R. et ANDERSON, S. R. (1992). *A-morphous morphology*. Numéro 62. Cambridge University Press.
- ANTONIOVA, V. et ŠTEKAUER, P. (2016). Derivational paradigms within selected conceptual fields—contrastive research. *Facta Universitatis, Series : Linguistics and Literature*, pages 61–75.
- APOTHÉLOZ, D. (2007). La préfixation en re-, l'antonymie directionnelle et les phénomènes de polarité sémantique. *Journal of French Language Studies*, 17(2):143–158.
- ARONOFF, M. (1993). *Morphology by itself : Stems and inflectional classes*, volume 22. MIT press.
- BAKER, C. F. (2009). La sémantique des cadres et le projet framenet : une approche différente de la notion de «valence». *Langages*, (4):32–49.
- BAKER, C. F. (2017). Framenet : Frame semantic annotation in practice. *In Handbook of Linguistic Annotation*, pages 771–811. Springer.

- BAKER, C. F., FILLMORE, C. J. et LOWE, J. B. (1998). The berkeley framenet project. *In 36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics, Volume 1*, pages 86–90.
- BARONI, M., BERNARDINI, S., FERRARESI, A. et ZANCHETTA, E. (2009). The wacky wide web : a collection of very large linguistically processed web-crawled corpora. *Language resources and evaluation*, 43(3):209–226.
- BARSALOU, L. W. (2012). Frames, concepts, and conceptual fields. *In Frames, fields, and contrasts*, pages 21–74. Routledge.
- BAUER, L. (1983). *English word-formation*. Cambridge university press.
- BAUER, L. (1997). Derivational paradigms. *In Yearbook of morphology 1996*, pages 243–256. Springer.
- BAUER, L. (2019). Notions of paradigm and their value in word-formation. *Word Structure*, 12(2):153–175.
- BENDER, E. M., GEBRU, T., McMILLAN-MAJOR, A. et SHMITCHELL, S. (2021). On the dangers of stochastic parrots : Can language models be too big? *In Proceedings of the 2021 ACM conference on fairness, accountability, and transparency*, pages 610–623.
- BENIAMINE, S. (2018). *Classifications flexionnelles. Étude quantitative des structures de paradigmes*. Thèse de doctorat, Université Sorbonne Paris Cité-Université Paris Diderot (Paris 7).
- BLANK, A. (2003). Words and concepts in time : Towards diachronic cognitive onomasiology. *Trends In Linguistics Studies And Monographs*, 143:37–66.
- BLEVINS, J. P. (2006). Word-based morphology. *Journal of linguistics*, 42(3):531–573.
- BLEVINS, J. P. (2016). *Word and paradigm morphology*. Oxford University Press.

- BOAS, H. C. (2005). From theory to practice : Frame semantics and the design of framenet. *Semantik im lexikon*, 479:129–159.
- BOCHNER, H. (1993). *Simplicity in generative morphology*. De Gruyter Mouton.
- BOLOZKY, S. (1999). *Measuring productivity in word formation : The case of Israeli Hebrew*, volume 27. Brill.
- BONAMI, O. (2014). *La structure fine des paradigmes de flexion*. Thèse de doctorat, Université Paris 7-Denis Diderot.
- BONAMI, O. (2017). Predictability in inflection and word formation. *In Paper delivered at ParadigMo 2017 : First Workshop on Paradigmatic Word-formation Modeling, Toulouse*, volume 19, page 20.
- BONAMI, O. et BENIAMINE, S. (2016). Joint predictiveness in inflectional paradigms. *Word Structure*, 9(2):156–182.
- BONAMI, O. et BOYÉ, G. (2019). Paradigm uniformity and the french gender system. *Perspectives on morphology : Papers in honour of Greville G. Corbett*, pages 171–192.
- BONAMI, O., BOYÉ, G., DAL, G., GIRAUDO, H. et NAMER, F. (2018). *The lexeme in descriptive and theoretical morphology*. Language Science Press.
- BONAMI, O. et STRNADOVÁ, J. (2016). Derivational paradigms : pushing the analogy. *In A paper presented at the SLE conference, Naples, Italy, August-3 September*.
- BONAMI, O. et STRNADOVÁ, J. (2019). Paradigm structure and predictability in derivational morphology. *Morphology*, 29(2):167–197.
- BOOIJ, G. (1986). Form and meaning in morphology the case of dutch «agent nouns». *Linguistics*, 24(283):503–517.

- BOUIJ, G. (2010). Compound construction : Schemas or analogy? a construction morphology perspective. *Cross-disciplinary issues in compounding*, pages 93–108.
- BORIN, L., DANNÉLLS, D. et HEPPIN, K. F. (2021). *The Swedish FrameNet+*. John Benjamins Publishing Company.
- BOYÉ, G. (2011). Régularités et classes flexionnelles dans la conjugaison du français.
- BOYÉ, G. et SCHALCHLI, G. (2017). Realistic paradigm for derivational morphology. *In ParadigMo 2017, first workshop on paradigmatic word formation modeling*, pages 19–20.
- BOYÉ, G. et SCHALCHLI, G. (2019). Realistic data and paradigms : The paradigm cell finding problem. *Morphology*, 29(2):199–248.
- BROWN, T., MANN, B., RYDER, N., SUBBIAH, M., KAPLAN, J. D., DHARIWAL, P., NEELAKANTAN, A., SHYAM, P., SASTRY, G., ASKELL, A. *et al.* (2020). Language models are few-shot learners. *Advances in neural information processing systems*, 33:1877–1901.
- BUSA, F. (1997). The semantics of agentive nominals in the generative lexicon. *Predicative Forms in Natural Language*.
- BUSA, F., CALZOLARI, N., LENCI, A. et PUSTEJOVSKY, J. (2001). Building a semantic lexicon : Structuring and generating concepts. *Computing Meaning : Volume 2*, pages 29–51.
- CANDITO, M., AMSILI, P., BARQUE, L., BENAMARA, F., DE CHALENDAR, G., DJEMAA, M., HAAS, P., HUYGHE, R., MATHIEU, Y. Y., MULLER, P. *et al.* (2014). Developing a french framenet : Methodology and first results. *In LREC-The 9th edition of the Language Resources and Evaluation Conference*.
- CARLSON, G. N. (1977). *Reference to kinds in English*. University of Massachusetts Amherst.
- CARSTAIRS-MCCARTHY, A. (1991). Inflection classes : Two questions with one answer. *Paradigms : The economy of inflection*, pages 213–53.

- CARSTAIRS-MCCARTHY, A. (1994). Inflection classes, gender, and the principle of contrast. *Language*, pages 737–788.
- CASSIDY, P. (2000). An investigation of the semantic relations in the roget’s thesaurus : Preliminary results. In *CICLing-2000 : Conference on Intelligent Text Processing and Computational Linguistics, February*, pages 13–19. Citeseer.
- CORBETT, G. G. (2010). Canonical derivational morphology. *Word structure*, 3(2):141–155.
- CROFT, W. (1991). *Syntactic categories and grammatical relations : The cognitive organization of information*. University of Chicago Press.
- CROFT, W., CRUSE, D. A. *et al.* (2004). *Cognitive linguistics*. Cambridge University Press.
- CRUSE, D. A. (2002). Language, meaning and sense : semantics. In *An encyclopedia of language*, pages 87–104. Routledge.
- CRUSE, D. A., CRUSE, D. A., CRUSE, D. A. *et* CRUSE, D. A. (1986). *Lexical semantics*. Cambridge university press.
- DAL, G. (2007). Les adverbes de manière en-ment du français : Dérivation ou flexion. N. Hathout *et* F. Montermini, *éds.*, *Morphologie à Toulouse*, (37):121–147.
- DAL, G. (2018). Les adverbes en-ment du français : Lexèmes ou formes d’adjectifs ? *The lexeme in descriptive and theoretical morphology*, page 87.
- DAL, G., HATHOUT, N., LIGNON, S., NAMER, F. *et* TANGUY, L. (2018). Toile versus dictionnaires : Les nominalisations du français en-age *et* en-ment. In *SHS Web of Conferences*, volume 46, page 08003. EDP Sciences.
- DAL, G. *et* NAMER, F. (2016). À propos des occasionnalismes. Paris, EDP Sciences, SHS Web of Conferences 27.

- DANLOS, L., PRADET, Q., BARQUE, L., NAKAMURA, T. et CONSTANT, M. (2016). Un verbenet du français. *Revue TAL*, 57(1):25.
- DAVIDSON, D. (1967). The logical form of action sentences. *The logic of decision and action*.
- DAVIDSON, D. (1969). *The individuation of events*. Springer.
- DJEMAA, M. (2017). *Stratégie domaine par domaine pour la création d'un FrameNet du français : annotations en corpus de cadres et rôles sémantiques*. Thèse de doctorat, Université Sorbonne Paris Cité.
- DOLBEY, A., ELLSWORTH, M. et SCHEFFCZYK, J. (2006). Bioframenet : A domain-specific framenet extension with links to biomedical ontologies. *In KR-MED*, volume 222.
- DOWTY, D. (1991). Thematic proto-roles and argument selection. *language*, 67(3):547–619.
- DRESSLER, W. U. (1989). Prototypical differences between inflection and derivation. *STUF-Language Typology and Universals*, 42(1):3–10.
- DUBOIS, J. et DUBOIS-CHARLIER, F. (1996). Collectifs d'êtres vivants. *Linx*, 34(1):125–132.
- EVANS, V. (2006). *Cognitive linguistics*. Edinburgh University Press.
- FELLBAUM, C. (2010). Wordnet. *In Theory and applications of ontology : computer applications*, pages 231–243. Springer.
- FERNÁNDEZ-DOMÍNGUEZ, J. (2019). The onomasiological approach. *In Oxford Research Encyclopedia of Linguistics*.
- FERNÁNDEZ-DOMÍNGUEZ, J., BAGASHEVA, A. et CLARES, C. L. (2020). *Paradigmatic Relations in Word Formation*. Brill.
- FILLMORE, C. (1968). The case for case. *Universals in Linguistic Theory*.
- FILLMORE, C. J. (1967). The case for case.

- FILLMORE, C. J. (1977). The case for case reopened. *In Grammatical relations*, pages 59–81. Brill.
- FILLMORE, C. J. (1985). Frames and the semantics of understanding. *Quaderni di semantica*, 6(2):222–254.
- FILLMORE, C. J. (2006). Frame semantics. *Cognitive linguistics : Basic readings*, 34:373–400.
- FILLMORE, C. J. *et al.* (1976). Frame semantics and the nature of language. *In Annals of the New York Academy of Sciences : Conference on the origin and development of language and speech*, volume 280, pages 20–32. New York.
- FILLMORE, C. J. *et* ATKINS, B. T. (1992). Toward a frame-based lexicon : The semantics of risk and its neighbors. *Frames, fields and contrasts : New essays in semantic and lexical organization*, 75:102.
- FILLMORE, C. J. *et* BAKER, C. F. (2001). Frame semantics for text understanding. *In Proceedings of WordNet and Other Lexical Resources Workshop, NAACL*, volume 6.
- FLAUX, N. *et* Van de VELDE, D. (2000). *Les noms en français : esquisse de classement*. Editions Ophrys.
- FRADIN, B. (2003). *Nouvelles approches en morphologie*. Presses universitaires de France.
- FRADIN, B. (2007). On the semantics of denominal adjectives. *In Mediterranean Morphology Meetings*, volume 6, pages 84–98.
- FRADIN, B. (2018). Paradigms and the role of series in derivational morphology. *Lingue e linguaggio*, 17(2):155–172.
- FRADIN, B. (2019). Competition in derivation : what can we learn from french doublets in-age and-ment ? *Competition in inflection and word-formation*, pages 67–93.

- FRADIN, B. (2020a). Characterizing derivational paradigms. *In Paradigmatic relations in word formation*, pages 49–84. Brill.
- FRADIN, B. (2020b). On meaning predictability in derivational domains. *In IMM19, 6-8 February 2020. Vienna : Wirtschaftsuniversität Wien.*
- FRADIN, B. (2021). Caractériser les paradigmes dérivationnels. *Verbum, XLIII*, 1:149–178.
- FRADIN, B. et WINTERSTEIN, G. (2012). Tuning agentivity and instrumentality : deverbal nouns in-oir revisited. *Paper delivered at Décembrettes*, 8:6–7.
- GILARDI, L. et BAKER, C. F. (2018). Learning to align across languages : Toward multilingual framenet. *In Proceedings of the international FrameNet workshop*, pages 13–22.
- GRUBER, J. S. (1965). *Studies in lexical relations*. Thèse de doctorat, Massachusetts Institute of Technology.
- HAAS, P., BARQUE, L., HUYGHE, R. et TRIBOUT, D. (2023). Pour une classification sémantique des noms en français appuyée sur des tests linguistiques. *Journal of French Language Studies*, 33(1):52–81.
- HARRIS, Z. S. (1954). Distributional structure. *Word*, 10(2-3):146–162.
- HASPELMATH, M. et SIMS, A. (2013). *Understanding morphology*. Routledge.
- HATHOUT, N. (2009). *Contributions à la description de la structure morphologique du lexique et à l'approche extensive en morphologie*. Thèse de doctorat, Université Toulouse le Mirail-Toulouse II.
- HATHOUT, N. (2011). Une approche topologique de la construction des mots : propositions théoriques et application à la préfixation en anti. *Des unités morphologiques au lexique*, pages 251–318.

- HATHOUT, N. et NAMER, F. (2014). Démonette, a french derivational morpho-semantic network. *In Linguistic Issues in Language Technology, Volume 11, 2014-Theoretical and Computational Morphology : New Trends and Synergies*.
- HATHOUT, N. et NAMER, F. (2018). La parasynthèse à travers les modèles : des rcl au parad. *The lexeme in descriptive and theoretical morphology, Language science Press, Berlin, Empirically Oriented Theoretical Morphology and Syntax*, pages 365–399.
- HATHOUT, N. et NAMER, F. (2019). Paradigms in word formation : what are we up to? *Morphology*, 29(2):153–165.
- HATHOUT, N. et NAMER, F. (2022). Paradis : a family and paradigm model. *Morphology*, pages 1–43.
- HATHOUT, N., NAMER, F. et ROCHÉ, M. (2021). French ethnonyms, toponyms, demonyms and their paradigmatic organization. *ParadigMo II*, page 33.
- HATHOUT, N., SAJOUS, F., CALDERONE, B. et NAMER, F. (2020). Glawinette : a linguistically motivated derivational description of french acquired from glawi. *In Proceedings of the 12th Language Resources and Evaluation Conference*, pages 3877–3885.
- HIGGINBOTHAM, J. (2000). On events in linguistic semantics. *Speaking of events*, 49:80.
- HUYGHE, R. (2015). Les typologies nominales : présentation. *Langue française*, 185(1):5–27.
- HUYGHE, R. (2021). La construction morphosémantique des noms d’agent en français. *Nommer l’humain. Descriptions, catégorisation, enjeux*, pages 189–204.
- HUYGHE, R. et TRIBOUT, D. (2015). Noms d’agents et noms d’instruments : le cas des déverbaux en-eur. *Langue française*, (1):99–112.
- HUYGHE, R. et VARVARA, R. (2023). Affix rivalry : Theoretical and methodological challenges. *Word Structure*, 16(1):1–23.

- HUYGHE, R. et WAUQUIER, M. (2020). What's in an agent ? a distributional semantics approach to agent nouns in french. *Morphology*, 30:185–218.
- JACKENDOFF, R. et AUDRING, J. (2019). Relational morphology in the parallel architecture. *In The Oxford handbook of morphological theory*.
- JURAFSKY, D. et MARTIN, J. H. (2014). Speech and language processing. vol. 3. *US : Prentice Hall*.
- KOCH, P. (1999). Frame and contiguity. *Metonymy in Language and Thought*. Amsterdam/-Philadelphia : John Benjamins Publishing Company, pages 139–167.
- KOCH, P. (2000). Pour une approche cognitive du changement sémantique lexical : aspect onomasiologique.
- KOCH, P. (2005). Aspects cognitifs d'une typologie lexicale synchronique. les hiérarchies conceptuelles en français et dans d'autres langues 1. *Langue française*, (01):11–33.
- KÖRTVÉLYESSY, L., BAGASHEVA, A., ŠTEKAUER, P., VALERA, S. et GENČI, J. (2020). 49 derivational networks in european languages : A cross-linguistic perspective. *In Derivational Networks Across Languages*, pages 485–608. De Gruyter Mouton.
- LAFOURCADE, M. et JOUBERT, A. (2008). Jeuxdemots : un prototype ludique pour l'émergence de relations entre termes. *In JADT'08 : Journées internationales d'Analyse statistiques des Données Textuelles*, pages 657–666.
- LAKOFF, G. (1988). Cognitive semantics. *Meaning and Mental Representation*.
- LANGACKER, R. W. (1987). *Foundations of cognitive grammar : Theoretical prerequisites*, volume 1. Stanford university press.
- LASSERRE, M. et MONTERMINI, F. (2014a). How is the meaning of complex lexemes constructed ? a study of neoclassical compounds in-cratie/-crate and-logie/-logue. *Italian Journal of Linguistics*, 26(2):157–181.

- LASSERRE, M. et MONTERMINI, F. (2014b). Pour une typologie des lexèmes construits : entre composition, composition néoclassique et affixation. *In SHS Web of Conferences*, volume 8, pages 1797–1812. EDP Sciences.
- LEHMANN, F. (1992). Semantic networks. *Computers & Mathematics with Applications*, 23(2-5):1–50.
- LEVIN, B. (1993). *English verb classes and alternations : A preliminary investigation*. University of Chicago press.
- LEVIN, B. et HOVAV, M. R. (2005). *Argument realization*. Cambridge University Press Cambridge.
- LIGNON, S. (2007). Les noms de spécialistes en-iste et en-ien : le chimiste perturbé ou comment le physicien se réajuste.
- LÖBNER, S. (2013). *Understanding semantics*. Routledge.
- LÖBNER, S. (2021). Frames at the interface of language and cognition. *Annual Review of Linguistics*, 7:261–284.
- LÖNNEKER-RODMAN, B. et BAKER, C. F. (2009). The framenet model and its applications. *Natural Language Engineering*, 15(3):415–453.
- LOWE, J. B. (1997). A frame-semantic approach to semantic annotation. *In Tagging Text with Lexical Semantics : Why, What, and How ?*
- LUCY, L. et BAMMAN, D. (2021). Gender and representation bias in gpt-3 generated stories. *In Proceedings of the Third Workshop on Narrative Understanding*, pages 48–55.
- LYONS, J. (1977). *Semantics : Volume 2*, volume 2. Cambridge university press.
- MAIENBORN, C. (2011). Event semantics. *Edited by Claudia Maienborn Klaus von Heusinger*, page 232.

- MATTHEWS, P. H. (1974). *Morphology*. Cambridge university press.
- MCNALLY, L. et BOLEDA, G. (2004). Relational adjectives as properties of kinds. Colloque de Syntaxe et Sémantique à Paris.
- MILLER, G. A. (1998). *WordNet : An electronic lexical database*. MIT press.
- MILLER, G. A., BECKWITH, R., FELLBAUM, C., GROSS, D. et MILLER, K. J. (1990). Introduction to wordnet : An on-line lexical database. *International journal of lexicography*, 3(4):235–244.
- MINSKY, M. (1974). *A framework for representing knowledge*. de Gruyter.
- MISSUD, A. et VILLOING, F. (2020). The morphology of rival-ion,-age and-ment selected verbal bases. *Lexique*, 26:29–52.
- MORRIS, J. et HIRST, G. (2004). Non-classical lexical semantic relations. *In Proceedings of the computational lexical semantics workshop at hlt-naacl 2004*, pages 46–51.
- MURPHY, M. L. (2003). *Semantic relations and the lexicon : Antonymy, synonymy and other paradigms*. Cambridge University Press.
- NAMER, F. (2009). *Morphologie, lexique et traitement automatique des langues*. Hermès-Lavoisier.
- NAMER, F. (2021). Adjectifs dénominaux, interprétation multiple et indices morpho-sémantiques. *Verbum (Presses Universitaires de Nancy)*, 63(1):3–12.
- NAMER, F. et HATHOUT, N. (2020). Paradis and démonette—from theory to resources for derivational paradigms. *The Prague Bulletin of Mathematical Linguistics*, 114(1):5–34.
- NAMER, F. et VILLOING, F. (2008). Interpréter les noms déverbaux : quelle relation avec la structure argumentale du verbe de base ? le cas des noms en oir (e). *In Congrès Mondial de Linguistique Française*, page 144. EDP Sciences.

- NAMER, F. et VILLOING, F. (2015). Sens morphologiquement construit et procédés concurrents : les noms de spécialistes en-logue et-logiste. *Revue de sémantique et pragmatique*, 35(35-36):7–26.
- NASTASE, V., SZPAKOWICZ, S., NAKOV, P. et SÉAGDHA, D. Ó. (2022). *Semantic relations between nominals*. Springer Nature.
- NUTTER, J. T. (1989). A lexical relation hierarchy. Rapport technique, Department of Computer Science, Virginia Polytechnic Institute & State
- OHARA, K. H., FUJII, S., OHORI, T., SUZUKI, R., SAITO, H. et ISHIZAKI, S. (2004). The japanese framenet project : An introduction. In *Proceedings of LREC-04 Satellite Workshop “Building Lexical Resources from Semantically Annotated Corpora”(LREC 2004)*, pages 9–11.
- PARSONS, T. (1995). Thematic relations and arguments. *Linguistic Inquiry*, pages 635–662.
- PELLEGRINI, M. (2020). Patterns of interpredictability and principal parts in latin verb paradigms : an entropy-based approach. *Journal of Latin Linguistics*, 19(2):195–229.
- PETERSEN, W. et GAMERSCHLAG, T. (2014). Why chocolate eggs can taste old but not oval : A frame-theoretic analysis of inferential evidentials. In *Frames and Concept Types*, pages 199–218. Springer.
- PETRUCK, M. R. (1996). Frame semantics. *Handbook of pragmatics*, 2.
- PETUKHOVA, V., BUNT, H. *et al.* (2008). Lyrics semantic role annotation : Design and evaluation of a set of data categories. In *LREC*. Citeseer.
- PLANK, F. (1994). Inflection and derivation. the encyclopedia of language and linguistics, vol. by re asher.
- PUSTEJOVSKY, J. (1998). *The generative lexicon*. MIT press.

- PUSTEJOVSKY, J. et JEZEK, E. (2016). Integrating generative lexicon and lexical semantic resources. *LREC 2016 Tutorial*.
- QUILLAN, M. R. (1966). Semantic memory. Rapport technique, Bolt Beranek and Newman Inc Cambridge MA.
- ROCHÉ, M. (2008). Structuration du lexique et principe d'économie : le cas des ethniques. *In Congrès mondial de linguistique française*, page 145. EDP Sciences.
- ROCHÉ, M. (2009). Pour une morphologie lexicale. *Mémoires de la Société de Linguistique de Paris*, (Nouvelle série n° 17):65–87.
- ROCHÉ, M. (2010). Base, thème, radical. *Recherches linguistiques de Vincennes*, (39):95–134.
- ROCHÉ, M. (2011). Pression lexicale et contraintes phonologiques dans la dérivation en-aie du français. *Linguistica*, 51(1):5–22.
- ROCHÉ, M. (2017a). Un exemple de réseau constructionnel : ethnique, toponymes, gentilés. *Toulouse : Université Toulouse*, 2.
- ROCHÉ, M. (2017b). Les familles dérivationnelles : comment ça marche. *Toulouse : Université Toulouse*, 2.
- ROCHÉ, M. et PLÉNAT, M. (2014). Le jeu des contraintes dans la sélection du thème présuffixal. *In SHS Web of Conferences*, volume 8, pages 1863–1878. EDP Sciences.
- ROCHÉ, M. et PLÉNAT, M. (2016). De l'harmonie dans la construction des mots français. *In SHS Web of Conferences*, volume 27, page 08003. EDP Sciences.
- ROSCH, E. H. (1973). Natural categories. *Cognitive psychology*, 4(3):328–350.
- ROSENBERG, M. (2008). *La formation agentive en français. Les composés [VN/A/Adv/P] N/A et les dérivés V-ant, V-eur et V-oir (e)*. Thèse de doctorat, Stockholm University ; Dept. of French, Italian and Classical Languages.

- RUPPENHOFER, J., ELLSWORTH, M., SCHWARZER-PETRUCK, M., JOHNSON, C. R. et SCHEFFCZYK, J. (2016). Framenet ii : Extended theory and practice.
- RUZ, A. E., FERNÁNDEZ-ALCAINA, C. et LARA-CLARES, C. (2022). *Paradigms in Word Formation : Theory and Applications*, volume 225. John Benjamins Publishing Company.
- SAGOT, B. et FIŠER, D. (2008). Building a free french wordnet from multilingual resources. *In OntoLex*.
- SALVADORI, J. et HUYGHE, R. (2023). Affix polyfunctionality in french deverbal nominalizations. *Morphology*, 33(1):1–39.
- SCHÄFER, R. (2015). Processing and querying large web corpora with the cow14 architecture. *In Proceedings of the 3rd Workshop on Challenges in the Management of Large Corpora*, pages 28–34.
- SCHÄFER, R. et BILDHAUER, F. (2012). Building large corpora from the web using a new efficient tool chain. *In Lrec*, pages 486–493.
- SCHALCHLI, G. et BOYÉ, G. (2018). Paradigms and syncretism in derivation : The case of ethnics in french. *Lingue e linguaggio*, 17(2):197–216.
- SCHANK, R. C. et ABELSON, R. P. (1975). Scripts, plans, and knowledge. *In IJCAI*, volume 75, pages 151–157.
- SCHMIDT, T. (2009). The kicktionary—a multilingual lexical resource of football language. *In Multilingual FrameNets in computational lexicography : methods and applications*, pages 101–132. de Gruyter.
- SCHNEDECKER, C. et ALEKSANDROVA, A. (2016). Les noms d’humains en-aire : essai de classification. *In 5ième congrès mondial de linguistique française*.
- SCHULER, K. K. (2005). *VerbNet : A broad-coverage, comprehensive verb lexicon*. University of Pennsylvania.

- SOWA, J. F. (1992). Semantic networks. *Encyclopedia of artificial intelligence*, 2:1493–1511.
- SOWA, J. F. (2014). *Principles of semantic networks : Explorations in the representation of knowledge*. Morgan Kaufmann.
- SPENCER, A. (2013). *Lexical relatedness*. OUP Oxford.
- ŠTEKAUER, P. (2014). Derivational paradigms. *The Oxford handbook of derivational morphology*, pages 354–369.
- STRNADOVÁ, J. (2014). *Les réseaux adjectivaux. Sur la grammaire des adjectifs dénominaux en français*. Thèse de doctorat, Université Paris Diderot (Paris 7) Sorbonne Paris Cité; Univerzita Karlova . . .
- STUMP, G. (2019). Some sources of apparent gaps in derivational paradigms. *Morphology*, 29(2):271–292.
- STUMP, G. T. (1991). A paradigm-based theory of morphosemantic mismatches. *Language*, pages 675–725.
- STUMP, G. T. (2005). Word-formation and inflectional morphology. *In Handbook of word-formation*, pages 49–71. Springer.
- SUBIRATS, C. et SATO, H. (2003). Surprise ! spanish framenet. *In In Proceedings of the Workshop on Frame Semantics at the XVII. International Congress of Linguists*. Citeseer.
- TALMY, L. (2000). *Toward a cognitive semantics*, volume 2. MIT press.
- TORRENT, T. T., BAKER, C. F., CZULO, O., OHARA, K. et PETRUCK, M. R. (2020). Proceedings of the international framenet workshop 2020 : Towards a global, multilingual framenet. *In Proceedings of the International FrameNet Workshop 2020 : Towards a Global, Multilingual FrameNet*.

- TORRENT, T. T., ELLSWORTH, M., BAKER, C. F. et MATOS, E. E. (2018). The multilingual fraMENET shared annotation task : a preliminary report. *In Proceedings of the Eleventh International Conference on Language Resources and Evaluation*, pages 62–68.
- TRIBOUT, D., LIGOZAT, A.-L. et BERNHARD, D. (2012). Constitution automatique d’une ressource morphologique : Verbagent. *In SHS Web of Conferences*, volume 1, pages 2517–2528. EDP Sciences.
- VAN MARLE, J. (1985). *On the paradigmatic dimension of morphological creativity*. De Gruyter.
- VENTURI, G., LENCI, A., MONTEMAGNI, S., VECCHI, E. M., SAGRI, M. T., TISCORNIA, D. et AGNOLONI, T. (2009). Towards a fraMENET resource for the legal domain. *In Proceedings of the 3rd Workshop on Legal Ontologies and Artificial Intelligence Techniques : 2nd Workshop on Semantic Processing of Legal Text*, pages 67–76.
- VILLOING, F. et NAMER, F. (2014). Composition néoclassique en-logue et-logiste : les noms en-logue sont-ils encore des noms de spécialistes. *Verbum*, 34(2):213–231.
- WALTHER, G. (2013). *Sur la canonicité en morphologie : Perspective empirique, formelle et computationnelle*. Thèse de doctorat, Université Paris 7 Denis Diderot.
- WINSTON, M. E., CHAFFIN, R. et HERRMANN, D. (1987). A taxonomy of part-whole relations. *Cognitive science*, 11(4):417–444.
- WUNDERLICH, D. et FABRI, R. (1995). Minimalist morphology : An approach to inflection. *Zeitschrift für Sprachwissenschaft*, 14(2):236–294.
- WURZEL, W. U. (1989). *Inflectional morphology and naturalness*, volume 9. Springer Science & Business Media.
- ZWANENBURG, W. (1988). Flexion et dérivation : le féminin en français. *Aspects de linguistique française. Hommage à QIM Mok*. Amsterdam : Rodopi, pages 191–208.

