



**Similarités et disparités sémantiques
entre verbes modaux du français :
Étude des propriétés combinatoires
de *pouvoir, devoir, falloir* et *vouloir* dans
une approche de linguistique de corpus**

Thèse présentée à la Faculté des Lettres et Sciences humaines
pour l'obtention du grade de doctoresse ès Lettres

par

Aylin Pamuksaç

Sous la direction de Prof. Corinne Rossari, Université de Neuchâtel, Suisse

Membres du jury :

Marion Bendinelli, Maitresse de Conférences, Université de Franche-Comté, France

Anouch Bourmayan, Maitresse de Conférences, Sorbonne Université, France

Francesca Dell'Oro, Professeuse assistante FNS, Université de Neuchâtel, Suisse

Thèse soutenue le 17 septembre 2024 à l'Université de Neuchâtel

IMPRIMATUR

La Faculté des lettres et sciences humaines de l'Université de Neuchâtel, sur les rapports de Mme Corinne Rossari, directrice de thèse, professeure, Université de Neuchâtel ; Mme Marion Bendinelli, maîtresse de conférences, Université de Franche-Comté, Besançon ; Mme Anouch Bourmayan, maîtresse de conférences, Sorbonne Université, Paris ; Mme Francesca Dell'Oro, professeure assistante, Université de Neuchâtel autorise l'impression de la thèse présentée par Mme Aylin Pamuksaç en laissant à l'autrice la responsabilité des opinions énoncées.

Neuchâtel, le 17 septembre 2024



Le doyen
Loris Petris

*A Luciän, et à toutes les personnes qui ont la force
d'être elles-mêmes.*

*You know you're good when you can even do it
With a broken heart*

Taylor Swift

Remerciements

J'aimerais commencer par remercier ma directrice de thèse, la professeuse Corinne Rossari pour ses encouragements durant ces quatre années de doctorat, qui m'ont permis de vivre la thèse et tous ses à-côtés de la meilleure des manières possibles.

Je remercie également les trois autres membres de mon jury pour leurs conseils et remarques bienveillantes qui m'ont aidée à améliorer au mieux ce manuscrit. Merci Marion Bendinelli pour tes nombreuses notes sur mon document (mais aussi pour les chouettes moments qu'on a partagés à Naples lors des JADT), merci Anouch Bourmaysan d'avoir apporté ton regard de pragmaticienne sur des données quantitatives, et merci Francesca Dell'Oro pour ton enthousiasme concernant mes recherches durant ces quatre années.

Ce doctorat n'aurait pas été le même sans les personnes que j'ai pu côtoyer à l'Université de Neuchâtel. Je pense, bien sûr, d'abord à l'équipe de linguistique française : merci Linda Sanvido d'avoir été la meilleure copine de thèse possible (et de croquetas !), merci Claudia Ricci pour ta bienveillance extraordinaire et merci Chloé Tahar pour ta rigueur dans le peu de travail qu'on a eu la chance de partager.

Je pense également à celles qui ne sont plus de la partie mais sans qui ce doctorat n'aurait pas eu la même saveur : merci Cyrielle Montrichard de m'avoir appris tout ce que je sais sur les outils quantitatifs mais aussi pour tous ces moments passés ensemble au bassin bleu, ou chez toi, ou dans la rue un mercredi soir, jusqu'à pas d'heure, (et aussi de m'avoir fait découvrir la cancoillotte). Merci Laura Aubry de m'avoir suivie dans mes envies de linguistique féministe et d'avoir été présente, au bureau, à Paris ou au Jura pendant cette dernière année de thèse. Merci aussi à Jessica Chessex et à Iveta Walther d'avoir été là lors de mes balbutiements dans le monde académique.

Merci à toutes les autres personnes de PAM7, avec, pour commencer, les collègues de l'étage : Loanne Janin (la meilleure des féministes que je connaisse), Melissa Juillet (tes cookies et ta motivation vont me manquer), Tiziana Kowalczyk (ta thèse va tellement révolutionner le monde), Samuel Dubois (ça fait toujours du bien de revoir quelqu'un du plateau), Sophia Fiedler (je suis encore admirative de ta soutenance de thèse) et Solène Belogi (car on se soutient entre meufs célibataires badass). Merci Paola Marongiu et Helena Bermúdez Sabel d'avoir été là dès le début. Merci aux secrétaires Florence Waelchli et Nathalie Millot de m'avoir permis de me décharger de quelques charges administratives en toute sérénité. Merci aux dialectologues de m'avoir intégrée dans votre équipe alors que je ne connais rien aux patois : Merci au professeur Mathieu Avanzi de m'avoir fait confiance pour ces derniers mois, merci

Céline Rumpf pour la raclette (la vraie !) et de me supporter jusqu'au bout de la nuit quand j'en ai besoin, et merci Yannick Emery pour le chouette travail qu'on a pu faire ensemble (notamment les micro-trott à Festineuch).

Merci à toutes les autres doctorant·es que j'ai pu côtoyer à Neuchâtel : merci Sandy Maillard pour tes encouragements et toutes les belles discussions qu'on a eues, merci Pauline Lavanchy pour ta motivation et ton implication pour l'ADFLESH, merci Hasti Noghrechi pour ton sourire à chaque fois qu'on se croise, ne serait-ce que brièvement, et merci Maialen Blázquez González, la seule autre vraie swiftie que je connaisse dans cette université. Merci Luciän (tu sais déjà tout).

Merci à toutes les doctorant·es des autres universités suisses que j'ai eu de la chance de rencontrer pendant les journées CUSO et autres évènements. On commence par les fribourgeois·es : Un grand merci à Justine Salvadori pour ces loooooongs messages vocaux qu'on adore, merci Alizée Lombard pour ta joie de vivre sans faille et merci Matthieu Monney pour toutes ces discussions lors des repas de midi *cusotiens*. A Berne, je remercie Baptiste Bersier pour son engouement contagieux. A Lausanne, je remercie mes collègues de *modalité* Ana Keck et Clotilde Robin pour nos riches échanges ainsi que Lena Möschler et Tristan Bornoze que j'aurais aimé connaître encore plus. A Genève, merci Sandrine Favre pour ce coup de cœur amical qu'on a pu vivre. Merci à toutes les candidat·es de MT180, édition 2023, ce qu'on a partagé était unique.

Merci à toutes les autres professeur·es de ces universités également : merci Richard Huyghes pour tes conseils et encouragements tout au long de ce doctorat, merci Sandrine Zufferey d'avoir participé à notre journée doctorale sur les questions de genres en linguistique et merci à Jérôme Jacquin pour tes observations que tu as portées sur ma recherche lors de cette conférence à Grenoble. Merci également à Alain Kamber pour nos échanges riches qu'on a pu partager même après que j'aie quitté l'ILCF. Merci aussi à Virginie Lethier pour tous ces moments partagés entre Neuchâtel et les conférences.

J'aimerais bien sûr aussi remercier toutes les personnes qui ne sont pas du monde académique et sans qui ces quatre ans de doctorat auraient été compliqués : merci mes parents de me soutenir même si vous ne comprenez pas toujours ce que je fais, merci Selin d'être la meilleure des sœurs qu'une sœur puisse avoir, merci mes ami·es de longue date Noémie, Léa, Zoé, Fanny, Nadia, Rebecca, Célia, Sarah, Fabien et ceux que j'ai rencontré·es récemment, Emilie, Athina, Odile, Félicien, et ceux que j'oublie.

Merci Baptiste de m'avoir toujours encouragée, d'avoir cru en moi et de continuer à le faire alors que nos chemins se sont séparés.

Merci enfin à Mélusine, plus qu'une amie, tu es une âme sœur et cette dernière année n'aurait vraiment pas été la même sans toi. Petite enveloppe et cœur avec les doigts.

Résumé

Cette thèse explore les verbes modaux en français – *pouvoir*, *devoir*, *falloir*, et *vouloir* – à travers une approche de linguistique de corpus. La recherche se concentre sur les propriétés sémantiques de ces verbes en étudiant les noyaux sémantiques, les effets de sens, les valeurs modales de ces verbes, tout comme leurs usages grammaticalisés. La méthodologie repose sur l'étude des contextes privilégiés, en se concentrant notamment sur les noms en position de sujet grammatical et les verbes à l'infinitif modalisés. Cette approche permet de révéler des propriétés sémantiques stables et des variations influencées par les genres et les champs textuels.

Les résultats montrent que chaque verbe modal possède des propriétés combinatoires spécifiques. Par exemple, *pouvoir* se distingue par une grande flexibilité sémantique, tandis que *devoir* est fortement associé à des contextes de nécessité et d'obligation. *Falloir*, quant à lui, est souvent utilisé comme marqueur de focalisation plutôt que comme véritable verbe modal. *Vouloir* présente des nuances importantes entre son usage comme verbe modal et ses emplois plus grammaticalisés.

L'utilisation d'analyses quantitatives, telles que le calcul des cooccurrences à l'aide de l'indice de spécificité, l'analyse factorielle des correspondances (AFC) ou encore les analyses arborées, a permis de mettre en lumière les similitudes et divergences entre ces verbes. Ces méthodes ont révélé des préférences combinatoires inédites, enrichissant notre compréhension des verbes modaux et mettant en évidence l'importance de revenir aux textes pour une interprétation qualitative des résultats. Cette thèse apporte ainsi des contributions significatives à la théorie des verbes modaux et démontre l'efficacité des méthodes de linguistique de corpus pour l'étude de la sémantique et des valeurs modales.

Mots-clés : Modalité ; Verbes modaux ; Sémantique ; Linguistique de corpus ; Textométrie.

Abstract

This thesis explores French modal verbs – *pouvoir*, *devoir*, *falloir*, and *vouloir* – through a corpus linguistics approach. The research focuses on the semantic properties of these verbs by studying their semantic cores, effects of meaning, modal values, as well as their grammaticalized uses. The methodology is based on the study of privileged contexts, particularly focusing on nouns in the grammatical subject position and infinitive verbs that are modalized. This approach reveals stable semantic properties and variations influenced by genres and textual fields.

The results show that each modal verb has specific combinatory properties. For example, *pouvoir* stands out for its significant semantic flexibility, whereas *devoir* is strongly associated with contexts of necessity and obligation. *Falloir*, on the other hand, is often used as a focalization marker rather than a true modal verb. *Vouloir* presents interesting nuances between its use as a modal verb and its more grammaticalized uses.

The use of quantitative analyses, such as the calculation of co-occurrences using the specificity index, correspondence factor analysis (CFA), and tree analyses, has highlighted the similarities and differences between these verbs. These methods have revealed unprecedented combinatory preferences, enriching our understanding of modal verbs and emphasizing the importance of returning to the texts for qualitative interpretation of the results. This thesis thus makes significant contributions to the theory of modal verbs and demonstrates the effectiveness of corpus linguistics methods for studying the semantics and modal values.

Keywords : Modality ; Modal Verbs ; Semantics ; Corpus Linguistics ; Textometry.

Table des matières

Table des figures.....	19
1. Introduction	23
1.1. Historique des études sur la modalité et importance de la subjectivité pour cette notion	23
1.2. Une approche sémantique et outillée des verbes modaux	25
2. Quels verbes modaux considérer ?	31
3. Articulation de la recherche	37
4. La sémantique des verbes modaux.....	39
4.1. La polysémie des verbes modaux : pas qu'une question de sémantique.....	40
4.1.1. La grammaticalisation des verbes modaux	45
4.2. Pouvoir.....	50
4.2.1. Le noyau sémantique de <i>pouvoir</i>	50
4.2.2. Les effets de sens de <i>pouvoir</i>	52
4.2.3. Les valeurs modales de <i>pouvoir</i>	53
4.2.4. Attribution de l'effet de sens ou de la valeur modale de <i>pouvoir</i> à l'aide du co(n)texte	55
4.2.5. L'emploi post-modal de <i>pouvoir</i>	57
4.2.6. Conclusions sur <i>pouvoir</i>	59
4.3. Devoir.....	60
4.3.1. Le noyau sémantique de <i>devoir</i>	60
4.3.2. Les valeurs modales de <i>devoir</i>	61
4.3.3. Les effets de sens de <i>devoir</i>	62
4.3.4. Une interprétation aisée des propriétés sémantiques de <i>devoir</i>	63
4.3.5. La grammaticalisation de <i>devoir</i>	64
4.3.6. Interprétation des effets de sens et des valeurs modales de <i>devoir</i> grâce au sujet grammatical	64
4.3.7. Interprétation des effets de sens et des valeurs modales de <i>devoir</i> grâce à la complémentation verbale sous forme d'un verbe à l'infinitif	66
4.3.8. Interprétation des effets de sens et valeurs modales de <i>devoir</i> selon d'autres paramètres.....	67
4.3.8.1. <i>La négation</i>	68
4.3.8.2. <i>Le temps verbal</i>	68
4.3.9. Les méthodologies utilisées pour étudier <i>devoir</i>	69
4.3.10. Conclusions sur <i>devoir</i>	70
4.4. Falloir	71
4.4.1. L'étymologie et l'évolution du noyau sémantique de <i>falloir</i> à travers le temps	71
4.4.2. Le noyau sémantique de <i>falloir</i>	73
4.4.3. Les effets de sens de <i>falloir</i>	73
4.4.4. Les valeurs modales de <i>falloir</i>	76
4.4.5. La grammaticalisation ou désémantisation de <i>falloir</i>	77
4.4.5.1. <i>Quelques remarques générales</i>	77
4.4.5.2. <i>Faut croire</i>	78
4.4.5.3. <i>Faut dire</i>	80
4.4.6. Les méthodologies employées pour appréhender les propriétés sémantiques de <i>falloir</i>	82
4.4.6.1. <i>La traduction</i>	82
4.4.6.2. <i>Les outils de la linguistique de corpus</i>	83
4.4.7. Conclusions sur <i>falloir</i>	84
4.5. Vouloir	85
4.5.1. Pourquoi <i>vouloir</i> ne serait pas un verbe modal.....	85
4.5.2. Pourquoi <i>vouloir</i> est un verbe modal	88
4.5.3. Les valeurs modales de <i>vouloir</i>	89
4.5.4. Les effets de sens de <i>vouloir</i>	91
4.5.5. Le noyau sémantique de <i>vouloir</i>	93
4.5.6. La grammaticalisation de <i>vouloir</i>	94
4.5.7. Les méthodologies employées pour appréhender les propriétés sémantiques de <i>vouloir</i>	98

4.5.8. Conclusions sur <i>vouloir</i>	99
4.6. Comparaisons entre les verbes modaux	99
4.6.1. Les comparaisons entre <i>pouvoir</i> et <i>devoir</i>	100
4.6.2. Les comparaisons entre <i>devoir</i> et <i>falloir</i>	101
4.6.3. Les comparaisons entre <i>pouvoir</i> , <i>devoir</i> et <i>vouloir</i>	103
4.6.4. Les comparaisons entre <i>pouvoir</i> , <i>devoir</i> et <i>falloir</i>	105
Synthèse sur la sémantique des verbes modaux	107
5. Objectifs de la présente thèse	109
6. Méthodologie	113
6.1. La linguistique de corpus et la textométrie	113
6.2. Une approche corpus-based ou/et corpus driven	117
6.3. Les diverses notions mobilisées dans cette recherche	118
6.3.1. Les corpus	119
6.3.2. La fréquence	121
6.3.3. La cooccurrence	122
6.3.4. L'empan	123
6.4. Les calculs et représentations des résultats	123
6.4.1. L'indice de spécificité	124
6.4.2. L'analyse factorielle des correspondances	125
6.4.3. L'analyse arborée	129
6.5. Les outils et corpus employés dans cette thèse	130
6.5.1. Présentation des corpus	130
6.5.2. Les logiciels	134
6.5.2.1. TXM	134
6.5.2.2. Hyperbase	140
Synthèse sur la méthodologie	140
7. Résultats et analyses	143
7.1. Les noms en position sujet	146
7.1.1. La typologie des noms de Salvadori et al. (2021)	149
7.1.2. Préambule à la recherche	153
7.1.3. Les résultats dans LM	154
7.1.3.1. La granularité épaisse pour les formes lemmatisées	154
7.1.3.2. La granularité moyenne pour les formes lemmatisées	155
7.1.3.3. Les sujets de type événement de pouvoir et devoir en forme de lemme	157
7.1.3.4. Les sujets de type animé de pouvoir et de devoir pour les formes lemmatisées	160
7.1.3.5. Les sujets de type animé et institution avec vouloir en forme de lemme	163
7.1.3.6. La granularité épaisse pour les formes à l'indicatif présent	166
7.1.3.7. La granularité moyenne pour les formes à l'indicatif présent	167
7.1.3.8. Les sujets de type animé pour les formes à l'indicatif présent : le cas de personne	168
7.1.3.9. Les sujets de type événement de devoir à l'indicatif présent	171
7.1.3.10. La granularité épaisse pour les formes au conditionnel présent	173
7.1.3.11. La granularité moyenne pour les formes au conditionnel présent	174
7.1.3.12. Les sujets de type objet cognitif de vouloir au conditionnel présent	175
7.1.3.13. Les sujets de type événement de pouvoir et devoir au conditionnel présent	177
7.1.4. Les résultats dans WIKI	179
7.1.4.1. La granularité épaisse pour les formes lemmatisées	180
7.1.4.2. La granularité moyenne pour les formes lemmatisées	181
7.1.4.3. Les sujets de type objet naturel et cognitif de pouvoir sous sa forme lemmatisée	181
7.1.4.4. Les sujets de type objet cognitif de vouloir sous sa forme lemmatisée	183
7.1.4.5. Les sujets de type animé de devoir sous sa forme lemmatisée	186
7.1.4.6. La granularité épaisse pour les formes au présent de l'indicatif	188
7.1.4.7. La granularité moyenne pour les formes à l'indicatif présent	189
7.1.4.8. La granularité épaisse pour les formes au conditionnel présent	190
7.1.4.9. La granularité moyenne pour les formes au conditionnel présent	191

7.1.4.10. Les sujets de type objet naturel et cognitif de pouvoir et devoir au conditionnel présent.....	192
7.1.4.11. Les sujets de type objet cognitif de vouloir au conditionnel présent.....	196
7.1.5. Le cas de falloir	198
7.1.6. L' AFC et l'analyse arborée appliquée aux sujets nominaux	204
7.1.6.1. Les résultats dans LM pour les formes lemmatisées	205
7.1.6.2. Les résultats dans LM pour les formes au présent de l'indicatif et au conditionnel présent	210
7.1.6.3. Les résultats dans WIKI pour les formes lemmatisées	213
7.1.6.4. Les résultats dans WIKI pour les formes au présent de l'indicatif et au conditionnel présent ..	215
Synthèse sur les associations entre les verbes modaux et les sujets nominaux.....	218
7.2. La complémentation verbale sous forme de verbe à l'infinitif	220
7.2.1. Le dictionnaire <i>Les Verbes Français</i> de Dubois et Dubois-Charlier (1997).....	221
7.2.2. Préambule à la recherche	225
7.2.3. Les résultats dans LM	226
7.2.3.1. Les répartitions des verbes en cooccurrences significatives pour les formes lemmatisées.....	226
7.2.3.2. Les complémentations verbales de type verbes de communication.....	228
7.2.3.3. Les complémentations verbales de type verbe psychologique	235
7.2.3.4. Les répartitions des verbes en cooccurrences significatives pour les formes au présent de l'indicatif.....	238
7.2.3.5. Les complémentations verbales de type verbes psychologiques avec pouvoir et falloir au présent de l'indicatif.....	240
7.2.3.6. Les complémentations verbales de type verbes réalisation/mise en état et locatifs avec devoir au présent de l'indicatif.....	244
7.2.3.7. Les complémentations verbales de type verbes de communication avec vouloir au présent de l'indicatif.....	245
7.2.3.8. Les répartitions des verbes en cooccurrences significatives pour les formes au conditionnel présent	247
7.2.3.9. Les complémentations verbales de type verbe de communication et psychologiques avec pouvoir et vouloir au conditionnel présent.....	249
7.2.3.10. Les complémentations verbales de type verbes de entrée/sortie, don/privation et communication avec devoir au conditionnel présent.....	252
7.2.3.11. Les complémentations verbales de type verbes psychologiques avec falloir au conditionnel présent	255
7.2.4. Les résultats dans WIKI	256
7.2.4.1. Les répartitions des verbes en cooccurrences significatives pour les formes lemmatisées.....	257
7.2.4.2. Les complémentations verbales de type verbes de communication et psychologiques avec pouvoir pour ses formes lemmatisées	258
7.2.4.3. Les complémentations verbales de type verbes de don/privation et réalisation/mise en état avec devoir pour ses formes lemmatisées.....	260
7.2.4.4. Les complémentations verbales de type verbes psychologiques avec falloir pour ses formes lemmatisées.....	262
7.2.4.5. Les complémentations verbales de type verbes de communication avec vouloir pour ses formes lemmatisées.....	264
7.2.4.6. Les répartitions des verbes en cooccurrences significatives pour les formes au présent de l'indicatif.....	266
7.2.4.7. Les répartitions des verbes en cooccurrences significatives pour les formes au conditionnel présent	267
7.2.4.8. Les complémentations verbales de type verbes de communication pour pouvoir au conditionnel présent	269
7.2.4.9. Les complémentations verbales de type verbes de serrer/saisir/posséder et don/privation pour devoir au conditionnel présent.....	271
7.2.4.10. Les complémentations verbales de type verbes de union/réunion pour falloir au conditionnel présent	272
7.2.4.11. Les complémentations verbales de type verbes de communication pour vouloir au conditionnel présent	273
7.2.5. L' AFC et l'analyse arborée appliquée aux complémentations verbales sous forme de verbe à l'infinitif.....	274
7.2.5.1. Les résultats dans LM pour les formes lemmatisées	274
7.2.5.2. Les résultats dans LM pour les formes au présent de l'indicatif et au conditionnel présent	277
7.2.5.3. Les résultats dans WIKI pour les formes lemmatisées	280

7.2.5.4. Les résultats dans WIKI pour les formes au présent de l'indicatif et au conditionnel présent	283
Synthèse sur les associations entre les verbes modaux et leur complémentation verbale	285
7.3. Synthèse de l'analyse des cotextes des verbes modaux	290
7.3.1. Intérêt de cette analyse	290
7.3.2. Les résultats à retenir	292
7.3.2.1. Les propriétés qui ont été confirmées	292
7.3.2.2. Les propriétés inédites des verbes modaux	292
7.3.2.3. Devoir	293
7.3.2.4. Pouvoir	294
7.3.2.5. Falloir	295
7.3.2.6. Vouloir	295
7.3.3. Similarités et disparités entre les verbes modaux	296
7.3.4. Limites de cette étude	298
7.3.5. Perspectives de recherche	298
8. Conclusions et perspectives générales de la thèse	301
8.1. Rappel des objectifs de cette thèse	301
8.2. Principaux résultats de cette thèse	302
8.2.1. Vérification des hypothèses de recherche	302
8.2.2. Retour sur quelques résultats importants de cette thèse	304
8.2.2.1. Pouvoir est le plus flou des verbes modaux	304
8.2.2.2. Devoir : le plus modal des modaux ?	304
8.2.2.3. Falloir : un marqueur de focalisation plutôt qu'un verbe modal ?	305
8.2.2.4. Vouloir : entre figement et expression de son NS	306
8.2.2.5. Les ressemblances et divergences notables entre ces quatre verbes modaux	306
8.2.3. Quelques remarques générales sur l'analyse	307
8.3. Apports de la thèse	308
8.3.1. Implications des réflexions théoriques sur les verbes modaux	308
8.3.2. Implications des méthodes quantitatives pour l'étude des verbes modaux	309
8.4. Ouvertures et perspectives de recherche	310
8.4.1. Étude d'autres facteurs pour les caractéristiques sémantiques des verbes modaux	310
8.4.2. Étude de la grammaticalisation des verbes modaux	311
8.4.3. Annotations dans les corpus	311
8.5. Pour conclure	312
9. Bibliographie	315

Table des figures

Figure 1 : Les verbes modaux considérés dans les recherches qui traitent plus de deux verbes modaux.	33
Figure 2 : Explication du passage de la lecture du « manque » à la lecture de la valeur modale déontique de falloir dans Herslund (2003)	72
Figure 3 : Carte des liens entre les divers effets de sens de pouvoir, devoir et vouloir, selon Desclés (2003).....	104
Figure 4 : AFC des paraphrases de pouvoir dans Caron et Caron-Pargue (2003 : 166).....	105
Figure 5 : Niveaux de classification de nos corpus inspirés de Malrieu et Rastier (2001 : 549).....	132
Figure 6 : Capture d'écran de la fonction « propriétés » dans TXM avec le corpus LM.	135
Figure 7 : Capture d'écran de la fonction « lexicale » dans TXM avec le corpus WIKI.	136
Figure 8 : Capture d'écran de la fonction « index » dans TXM avec le corpus LM.	137
Figure 9 : Capture d'écran de la fonction « concordances » dans TXM avec le corpus WIKI.	138
Figure 10 : Capture d'écran de la fonction « cooccurrences » dans TXM avec le corpus WIKI	139
Figure 11 : Arbre de décision pour la classification des noms (Haas et al. 2022 : 23).....	151
Figure 14 : Répartition des noms en position sujet selon la granularité de Varvara et al. (sous presse) dans le corpus LM.	155
Figure 15 : Répartition des noms en position sujet pour les formes lemmatisées des modaux selon les catégories de Salvadori et al. (2021) dans le corpus LM.....	156
Figure 16 : Les 30 sujets nominaux de type événement avec le plus grand indice de spécificité pour pouvoir et devoir dans le corpus LM.....	157
Figure 17 : Les 30 sujets nominaux de type animé avec le plus grand indice de spécificité pour pouvoir et devoir dans le corpus LM.....	161
Figure 18 : Les 20 sujets nominaux avec l'indice de Lafon (1980) de vouloir sous sa forme lemmatisée dans le corpus dans le corpus LM.	164
Figure 19 : Répartition des noms en position sujet pour les formes à l'indicatif présent des modaux selon la granularité de Varvara et al. (sous presse) dans le corpus LM.	167
Figure 20 : Répartition des noms en position sujet pour les formes à l'indicatif présent des modaux selon les catégories de Salvadori et al. (2021) dans le corpus LM.....	168
Figure 21 : Les 30 sujets nominaux de type animé avec le plus grand indice de spécificité pour pouvoir, devoir et vouloir au présent de l'indicatif dans le corpus LM.....	169
Figure 22 : Les 30 sujets nominaux de type événement avec le plus grand indice de spécificité pour devoir au présent de l'indicatif dans le corpus LM.....	172
Figure 23 : Répartition des noms en position sujet pour les formes au conditionnel présent des modaux selon la granularité de Varvara et al. (sous presse) dans le corpus LM.	173
Figure 24 : Répartition des noms en position sujet pour les formes au conditionnel présent des modaux selon les catégories de Salvadori et al. (2021) dans le corpus LM.....	174
Figure 25 : Les 4 sujets nominaux de type objet cognitif pour vouloir au conditionnel présent dans le corpus LM selon l'indice de spécificité.....	175
Figure 26 : Les 30 sujets nominaux de type événement avec le plus grand indice de spécificité pour pouvoir et devoir au conditionnel présent dans le corpus LM.....	178
Figure 27 : Répartition des noms en position sujet pour les formes lemmatisées des modaux selon la granularité de Varvara et al. (sous presse) dans le corpus LM.	180
Figure 28 : Répartition des noms en position sujet pour les formes lemmatisées des modaux selon les catégories de Salvadori et al. (2021) dans le corpus WIKI.	181
Figure 29 : Les 30 sujets nominaux de type objet naturel et cognitif avec le plus grand indice de spécificité pour pouvoir sous sa forme lemmatisée dans le corpus WIKI.	182

Figure 30 : Les 10 sujets nominaux de type objet cognitif pour vouloir sous sa forme lemmatisée dans le corpus WIKI selon l'indice de spécificité.....	184
Figure 31 : Les 30 sujets nominaux de type animés avec le plus grand indice de spécificité pour devoir sous sa forme lemmatisée dans le corpus WIKI.	187
Figure 32 : Répartition des noms en position sujet pour les formes à l'indicatif présent des modaux selon la granularité de Varvara et al. (sous presse) dans le corpus WIKI.....	189
Figure 33 : Répartition des noms en position sujet pour les formes au conditionnel présent des modaux selon les catégories de Salvadori et al. (2021) dans le corpus WIKI.	190
Figure 34 : Répartition des noms en position sujet pour les formes au présent de l'indicatif des modaux selon la granularité de Varvara et al. (sous presse) dans le corpus WIKI.....	191
Figure 35 : Répartition des noms en position sujet pour les formes au conditionnel présent des modaux selon les catégories de Salvadori et al. (2021) dans le corpus WIKI.	192
Figure 36 : Les 30 sujets nominaux de type objet naturel et cognitif avec le plus grand indice de spécificité pour pouvoir et les 17 sujets nominaux de type objets naturels et cognitifs avec un indice de spécificité significatif pour devoir au conditionnel présent dans le corpus WIKI.	193
Figure 37 : Les 3 sujets nominaux de type objets naturels et cognitifs avec un indice de spécificité significatif pour vouloir au conditionnel présent dans le corpus WIKI.....	196
Figure 38 : Part des sujets nominaux exprimés dans la complétive pour falloir sous sa forme lemmatisée.	199
Figure 39 : Répartition des noms en position sujet selon la granularité de Varvara et al. (sous presse) pour falloir sous sa forme lemmatisée dans le corpus LM.	199
Figure 40 : Répartition des noms en position sujet pour les formes lemmatisées de falloir selon les catégories de Salvadori et al. (2021) dans le corpus WIKI.	200
Figure 41 : Les 24 sujets nominaux avec un indice de spécificité significatif pour falloir sous sa forme lemmatisée dans le corpus LM.	200
Figure 42 : AFC des modaux sous leurs formes lemmatisées avec les catégories sémantiques des noms en position de sujet grammatical (ou sémantique pour falloir) dans LM.	205
Figure 43 : Analyse arborée des modaux sous leurs formes lemmatisées avec les catégories sémantiques des noms en position de sujet grammatical (ou sémantique pour falloir) dans LM.	208
Figure 44 : AFC des verbes pouvoir, devoir et vouloir au présent de l'indicatif et au conditionnel présent avec les catégories sémantiques des noms en position de sujet grammatical dans LM.	210
Figure 45 : Analyse arborée des verbes pouvoir, devoir et vouloir au présent de l'indicatif et au conditionnel présent avec les catégories sémantiques des noms en position de sujet grammatical dans LM.	212
Figure 46 : AFC des modaux sous leurs formes lemmatisées avec les catégories sémantiques des noms en position de sujet grammatical (ou sémantique pour falloir) dans WIKI.	213
Figure 47 : Analyse arborée des modaux sous leurs formes lemmatisées avec les catégories sémantiques des noms en position de sujet grammatical (ou sémantique pour falloir) dans WIKI.	215
Figure 48 : AFC des verbes pouvoir, devoir et vouloir au présent de l'indicatif et au conditionnel présent avec les catégories sémantiques des noms en position de sujet grammatical dans WIKI.	216
Figure 49 : Analyse arborée des verbes pouvoir, devoir et vouloir au présent de l'indicatif et au conditionnel présent avec les catégories sémantiques des noms en position de sujet grammatical dans WIKI.	217
Figure 50 : Les 14 classes de Dubois et Dubois-Charlier (1997) présentées dans François et al. (2007).	224
Figure 51 : Entrées pour le verbe habiter dans le LVF électronique.	225
Figure 52 : Répartition des verbes à l'infinitif en position de complémentation verbale selon la classification de Dubois et Dubois-Charlier (1997) pour les formes lemmatisées des modaux dans le corpus LM.	227
Figure 53 : Indice de spécificité pour les verbes de communication en commun pour les quatre modaux sous leur forme lemmatisée dans le corpus LM.	228
Figure 54 : AFC des modaux sous leur forme lemmatisée avec les verbes de communication en commun en position de complémentation verbale dans les 30 indices les plus élevés dans LM.	234

Figure 55 : Indice de spécificité pour les verbes psychologiques en commun en position de complémentation verbale pour les quatre modaux étudiés sous leur forme lemmatisée dans le corpus LM.....	235
Figure 56 : Répartition des verbes à l'infinitif en position de complémentation verbale selon la classification de Dubois et Dubois-Charlier (1997) pour les formes du présent de l'indicatif des modaux dans le corpus LM.....	239
Figure 57 : Les 30 verbes de type psychologiques en position de complémentation verbale avec le plus grand indice de spécificité pour pouvoir et falloir au présent de l'indicatif dans LM.....	240
Figure 58 : Les 30 verbes de type locatif ou de réalisation/mise en état en position de complémentation verbale avec le plus grand indice de spécificité pour devoir au présent de l'indicatif dans LM.....	244
Figure 59 : Les 30 verbes de type communication en position de complémentation verbale avec le plus grand indice de spécificité pour vouloir au présent de l'indicatif dans LM.	245
Figure 60 : Répartition des verbes à l'infinitif en position de complémentation verbale selon la classification de Dubois et Dubois-Charlier (1997) pour les formes du conditionnel présent des modaux dans le corpus LM.	248
Figure 61 : Les (30) verbes de type communication et psychologique en position de complémentation verbale avec le plus grand indice de spécificité pour pouvoir et vouloir au conditionnel présent dans LM.....	250
Figure 62 : Les 30 verbes de type entrée/sortie, don/privation et communication en position de complémentation verbale avec le plus grand indice de spécificité pour devoir au conditionnel présent dans LM.....	253
Figure 63 : Les verbes de type psychologique en position de complémentation verbale avec le plus grand indice de spécificité pour falloir au conditionnel présent dans LM.	255
Figure 64 : Répartition des verbes à l'infinitif en position de complémentation verbale selon Dubois et Dubois-Charlier (1997) pour les formes lemmatisées des modaux dans le corpus WIKI.....	257
Figure 65 : Les 30 verbes de type communication et psychologiques en position de complémentation verbale avec le plus grand indice de spécificité pour pouvoir pour ses formes lemmatisées dans le corpus WIKI.....	258
Figure 66 : Les 30 verbes de type don/privation et réalisation/mise en état en position de complémentation verbale avec le plus grand indice de spécificité pour devoir pour ses formes lemmatisées dans le corpus WIKI.	261
Figure 67 : Les 30 verbes de type psychologique en position de complémentation verbale avec le plus grand indice de spécificité pour falloir pour ses formes lemmatisées dans le corpus WIKI.	262
Figure 68 : Les 30 verbes de communication en position de complémentation verbale avec le plus grand indice de spécificité pour vouloir pour ses formes lemmatisées dans le corpus WIKI.	264
Figure 69 : Répartition des verbes à l'infinitif en position de complémentation verbale selon la classification Dubois et Dubois-Charlier (1997) pour les formes du présent de l'indicatif des modaux dans le corpus WIKI.	266
Figure 70 : Répartition des verbes à l'infinitif en position de complémentation verbale selon la classification de Dubois et Dubois-Charlier (1997) pour les formes du conditionnel présent des modaux dans le corpus WIKI.	268
Figure 71 : Les 29 verbes de communication en position de complémentation verbale avec le plus grand indice de spécificité pour pouvoir au conditionnel présent dans le corpus WIKI.	269
Figure 72 : Les 22 verbes de type saisir/serrer/posséder et de don/privation en position de complémentation verbale avec le plus grand indice de spécificité pour devoir au conditionnel présent dans le corpus WIKI.	271
Figure 73 : AFC des verbes sous leurs formes lemmatisées avec les catégories sémantiques de Dubois et Dubois-Charlier (1997) des verbes en position complémentation verbale dans LM.	275
Figure 74 : Analyse arborée des modaux sous leurs formes lemmatisées avec les catégories sémantiques de Dubois et Dubois-Charlier (1997) des verbes en position de complémentation verbale dans LM.	277
Figure 75 : AFC des modaux au présent de l'indicatif et au conditionnel présent avec les catégories sémantiques de Dubois et Dubois-Charlier (1997) des verbes en position de complémentation verbale dans LM.....	278
Figure 76 : Analyse arborée des modaux au présent de l'indicatif et au conditionnel présent avec les catégories sémantiques des verbes de Dubois et Dubois-Charlier (1997) en position de complémentation verbale dans LM.	279
Figure 77 : AFC des verbes sous leurs formes lemmatisées avec les catégories sémantiques de Dubois et Dubois-Charlier (1997) des verbes en position complémentation verbale dans WIKI.	281

Figure 78 : Analyse arborée des verbes sous leurs formes lemmatisées avec les catégories sémantiques des verbes de Dubois et Dubois-Charlier (1997) en position de complémentation verbale dans WIKI.....	282
Figure 79 : AFC des modaux au présent de l'indicatif et au conditionnel présent avec les catégories sémantiques des verbes de Dubois et Dubois-Charlier (1997) en position de complémentation verbale dans WIKI.	283
Figure 80 : Analyse arborée des modaux au présent de l'indicatif et au conditionnel présent avec les catégories sémantiques des verbes de Dubois et Dubois-Charlier (1997) en position de complémentation verbale dans WIKI.	285

1. Introduction

La modalité est sans doute un des concepts les plus compliqués à caractériser en linguistique¹ car elle peut être définie d'autant de manières qu'il existe de recherches à ce propos : « La modalité et ses catégories peuvent être définies et nommées de nombreuses manières. Il n'y a pas de manière plus juste qu'une autre. La seule exigence est que la chercheuse² clarifie la manière dont elle utilise ces notions »³ (Van der Auwera & Plungian 1998 : 80). C'est cette tâche qui va nous occuper dans un premier temps avec un retour aux études et notions phares de ce concept de modalité.

1.1. Historique des études sur la modalité et importance de la subjectivité pour cette notion

Lorsqu'on traite la modalité en linguistique, il est nécessaire de faire un détour par la philosophie, car il n'existe, *a priori*, « [...] pas de textes plus anciens traitant le sujet [...] » (Le Querler 2004 : 644) que celui d'Aristote. D'ailleurs, cette notion a préoccupé pendant plusieurs siècles les philosophes uniquement et c'est bien plus tard que les linguistes se sont penchées sur le sujet (De Haan 1997 : 3).

Les distinctions qu'opérait alors Aristote étaient très dichotomiques : il existe des énoncés modaux, et d'autres qui ne le sont pas (Le Querler 2004 : 643-644). Les énoncés modaux sont ainsi constitués de « deux parties : ce dont on dit quelque chose (c'est-à-dire le sujet) et ce qui est dit (c'est-à-dire le prédicat). » (Chu 2008 : 16)⁴. La première composante pouvait dès lors arborer une valeur de *nécessité* ou de *possibilité*, aussi réduite aux notions de *vrai* et de *faux* selon certains interprètes d'Aristote (Larreya 2004 : Le Querler 2004).

Pour les études en linguistique, l'intérêt pour la modalité est surtout apparu au début des années 2000, comme le rappelle l'ouvrage de Narrog (2012c) dans lequel celui-ci écrit que la

¹ La modalité n'est pas seulement un sujet linguistique. Elle est aussi traitée dans d'autres domaines comme la philosophie et la logique (Larreya 2004 : 733).

² Nous faisons le choix, dans ce travail, d'utiliser le féminin générique pour les noms mais aussi pour les pronoms pour désigner les personnes. Ainsi, lorsqu'un groupe est composé de minimum une personne genrée au féminin, nous emploierons le pronom *elle-s*. Nous utiliserons le pronom *il-s* uniquement lorsque nous faisons référence à une personne ou un groupe de personnes qualifiées ou connues en tant qu'homme. Les noms et les pronoms restent cependant inchangés dans les citations d'autres autrices.

Dans le cas où nous faisons référence à d'autres éléments que des humains, nous adoptons l'accord de proximité.

³ Les citations issues de publications en anglais ou en allemand ont été traduites par nous-même afin de permettre une meilleure intégration au texte.

⁴ Ces deux notions peuvent se rapporter au *modus* et au *dictum* de Bally (1965) dont nous discuterons la théorie plus bas.

[...] modalité est devenue l'un des domaines les plus dynamiques des études linguistiques au cours de la dernière décennie, car les linguistes ont reconnu son importance pour le langage et la communication. Les êtres humains imaginent constamment des choses qui ne sont pas réelles mais qui restent possibles, voire contraires aux faits. (Narrog 2012b : 1)

Toutefois, comme le mentionnent Flaux et Lagae (2014), nous pouvons admettre que l'intérêt pour la modalité est apparu avec le courant de la pragmatique, et plus précisément avec les théories sur les actes de langage pour les études anglophones (Austin 1970 : Searle 1972, 1979), puis avec Ducrot (1984) et sa théorie sur la polyphonie pour les recherches francophones⁵. C'est donc par le biais des théories sur l'énonciation et l'intérêt pour le concept de *locutrice* que les études sur la modalité ont commencé à foisonner en linguistique française⁶ :

L'essor de la perspective énonciative a entraîné dans cette période un vif intérêt pour les modalités dans la mesure où celles-ci expriment l'attitude du locuteur par rapport aux événements qu'il relate et permettent donc de le saisir dans une activité presque aussi essentielle que celle qui a trait à la construction des repères déictiques de l'énonciation. On a d'ailleurs vu se répandre des expressions comme "sujet déictique" et "sujet modal" indiquant cette double activité du locuteur et laissant ouverte la possibilité d'une dissociation entre ces deux instances (cf. Rabatel 2005). (Monte 2011 : 85)

Cependant, avant cet intérêt grandissant pour la notion de modalité, Bally (1965) avait déjà esquissé une définition de ce qui pouvait être considéré comme faisant partie la modalité. La principale distinction que l'auteur a opérée est celle de *dictum* et de *modus*. Le premier est « la représentation reçue par les sens, la mémoire ou l'imagination » alors que le second est « l'opération psychique du sujet pensant » (Bally 1965 : 36). Le *modus* est ainsi un « jugement exprimé par un sujet modal (ou sujet pensant) au moyen d'un verbe modal » (Bendinelli 2012 : 32).

Traiter la question de la modalité est donc intrinsèquement traiter la question de la subjectivité dans le langage, ou, du moins, la distinction entre des énoncés objectifs ou subjectifs (Ducrot 1993). Cela « revient à inscrire la réflexion dans celle de l'étude des lieux d'expression de la subjectivité des sujets parlants » (Bendinelli 2012 : 25). Cependant, cette expression de la subjectivité peut être plus ou moins marquée, à tel point que Bendinelli (2012) parle d'un « continuum de manifestations du *modus*, mises en œuvre de stratégies énonciatives visant à dissimuler ou non la subjectivité du sujet » (Bendinelli 2012 : 38). Ce n'est donc pas anodin de

⁵ Avant cela, la modalité était surtout un concept morphosyntaxique qui permettait de traiter les différents types de phrases comme les déclaratives, les interrogatives, les injonctives, les exclamatives, etc. (Flaux & Lagae 2014 : 4)

⁶ Pour une chronologie plus détaillée sur l'histoire de la notion de modalité, se référer à Le Querler (2004) et Monte (2011).

noter que certaines théories sur la modalité, notamment celle de Culioli (1990, 1999a, 1999b), prennent en compte comme point de départ le degré d'engagement de la locutrice pour établir les diverses classes modales. Ainsi, dans ce modèle :

La modalité de rang I regroupe l'assertion (positive ou négative), l'interrogation et l'injonction : celle de rang II couvre un continuum de probabilités allant de l'expression de la simple éventualité à celle de la présomption d'inéluctabilité : celle de rang III inclut les jugements esthétiques, axiologiques, affectifs : enfin, la modalité de rang IV, à l'image de celle de rang II, couvre un continuum de valeurs sémantiques qui, en fonction du caractère animé ou non de l'agent responsable de l'actualisation de la relation prédicative, va de la permission à l'obligation (pour un agent animé uniquement), de la simple capacité à la nécessité ou au besoin, et englobe aussi leurs équivalents contradictoires respectifs. (Bendinelli 2012 : 49)

Pour notre part, nous n'adoptons pas ce type d'approche mais plutôt un raisonnement attachant davantage d'importance à la sémantique des classes modales. Dans cette thèse, nous nous intéresserons aux noyaux sémantiques, effets de sens et valeurs modales des verbes modaux *pouvoir*, *devoir*, *falloir* et *vouloir* dans une approche co(n)textuelle du sens. Ce co(n)texte sera étudié à l'aide des outils de la linguistique de corpus et de la textométrie⁷.

1.2. Une approche sémantique et outillée des verbes modaux

Comme énoncé plus haut, il peut être difficile de définir ce qu'est la modalité tant les approches s'étant intéressées à ce concept sont nombreuses. Cependant, dans une acception large de la modalité, il est admis qu'elle « [...] est l'expression de l'attitude du locuteur par rapport au contenu propositionnel de son énoncé » (Le Querler 2004 : 646). Toutefois, il est à noter que cette définition n'est pas suffisante car, à l'aide de verbes modaux, la locutrice peut également exprimer la capacité d'un individu à faire quelque chose (1) ou émettre une vérité absolue (2) :

1. Henry peut faire le rapport.
2. Si tu lances une pièce en l'air, elle **doit** retomber. (Exemple emprunté à Vetters (2021))

Dans les énoncés 1 et 2, nous constatons qu'il ne s'agit pas d'un point de vue de la locutrice sur un objet mais bel est bien d'une description d'un état de fait. C'est pourquoi il semble nécessaire de redéfinir la modalité d'une manière plus inclusive encore, en ajoutant, comme le fait Vetters (2021), que « [l]a *modalité* est l'attitude du locuteur vis-à-vis de la réalité ou l'existence d'une

⁷ Les outils seront détaillés dans le chapitre 6 de ce travail.

situation : **celui-ci peut en effet estimer que les choses ne peuvent pas être autrement, ou au contraire que la situation pourrait ou aurait pu ne pas exister**⁸ » (Vetters 2021 : 1349).

Ces deux exemples peuvent être classés dans la catégorie **des modalités descriptives** (Vetters 2021), aussi appelées **aléthiques** (Gosselin 2010 : Le Querler 1996) ou encore **dynamiques** (Nuyts 2016)⁹. Nous emploierons dans ce travail la terminologie de Vetters (2021), qui non seulement est la plus récente de notre travail, mais dont la définition est également la plus englobante étant donné que le chercheur définit ce type de modalités comme « les contraintes du monde et les capacités des individus » (Vetters 2021 : 1349)¹⁰. La dénomination *dynamique* (*dynamic*) est plutôt réservée aux études anglophones, à l'instar du *Oxford handbook of modality and mood* de Nuyts et Van Der Auwera (2016).

Cependant, cette catégorie n'est pas toujours considérée dans les recherches sur la modalité. En effet, pour certaines linguistes, ce sont les conceptions étroites de la modalité (ou restreintes comme le souligne Barbet (2013 : 4)), « principalement issues de la tradition logique (d'Aristote en particulier), [et] centrées sur les notions de nécessaire et de possible » (Gosselin 2010 : 6) qui font foi. La *nécessité* est exprimée à travers **les modalités déontiques** et la *possibilité* à travers **les modalités épistémiques**. Les premières, dans leur acception la plus sommaire, « [...] définissent ce qui doit être, conformément à une règle sociale, morale, etc. » (Le Querler 1996 : 41). Les secondes « marquent le domaine du certain, du douteux, du savoir, de la croyance » (Le Querler 1996 : 41) et manifestent ainsi le lien entre la locutrice et son degré de certitude concernant ce qu'elle énonce.

Ce n'est donc pas un hasard si ce sont principalement les verbes *devoir* et *pouvoir* qui sont les plus souvent analysés dans les études sur les marqueurs de la modalité : ils sont les prototypes mêmes de ces deux extrémités de l'échelle modale. Dell'Oro (accepté-a) nuance cependant ce propos en expliquant que c'est généralement dans les descriptions synchroniques sémantiques que nous retrouvons cette dichotomie entre le *nécessaire* et le *possible*. Dans d'autres perspectives, c'est davantage la notion de subjectivité et en particulier les modalités épistémiques qui sont étudiées.

⁸ Nous mettons en évidence.

⁹ Pour un panorama plus détaillé des équivalences terminologiques entre les autrices, se référer à Dell'Oro (accepté-a).

¹⁰ De plus, soulignons que, dans ses travaux plus anciens, Vetters utilise la dénomination *aléthique* (Vetters 2007, 2012 : Vetters & Barbet 2006) pour des usages similaires à ceux que nous citons ici et que ce terme n'apparaît plus dans sa publication la plus récente, soit celle de 2021. Les modalités aléthiques chez Le Querler (1996) et Gosselin (2010) sont également présentées avec des exemples similaires à ceux que nous avons mis en évidence ici, ce qui montre bien l'équivalence de ces termes.

Dans l'étude qui est la nôtre, nous prendrons en compte ces trois catégories modales car elles permettent de regrouper un grand nombre d'effets de sens liés aux verbes que nous allons étudier. Dans une moindre mesure, nous ajoutons également les **modalités bouliques**, qui elles, ne concernent que le verbe *vouloir* dans notre paradigme. Ces modalités « marquent la volonté du sujet » (Le Querler 1996 : 42). Cependant, il est intéressant de noter que, *stricto sensu*, dans une approche prenant en compte uniquement les modalités du nécessaire et du possible, les volontés peuvent traduire l'une des deux extrémités de cette échelle modale. Ainsi, Narrog (2012b) explique que la modalité boulique est « [une] proposition [...] marquée comme une nécessité ou une possibilité par rapport à la volonté ou aux intentions de quelqu'un » (Narrog 2012b : 9). Nous pouvons ainsi comprendre un énoncé comme l'exemple 3a comme une demande de faire (donc une modalité déontique en 3b) ou une possibilité (donc une modalité épistémique en 3c) :

3a) Je *veux* du chocolat.

3b) Il est *nécessaire* que je mange du chocolat, pour ne pas tomber dans les pommes

3c) J'aimerais avoir la *possibilité* de manger du chocolat, car j'en ai envie.

En plus des questions concernant les catégories modales à considérer, il est également pertinent de se demander quel type de marqueurs modaux nous pouvons observer. Dans les études sur la modalité, c'est généralement la catégorie grammaticale des verbes qui est analysée, et ce, dans plusieurs langues (Larrea 2004 ; Van der Auwera & Plungian 1998), au détriment d'autres éléments comme les noms (*la possibilité de...*), les adjectifs (*nécessaire*), ou encore les adverbes (*obligatoirement*) (Le Querler 2004 : 646).

L'importance des verbes dans le domaine modal peut être due à plusieurs autres facteurs. Pour commencer, les verbes ont la particularité de se décliner en un grand nombre de formes selon le mode¹¹ ou le temps verbal¹² auxquels ils sont conjugués, ce qui peut en changer l'effet de sens et donc la valeur modale¹³ qu'ils véhiculent (Vetters & Barbet 2006 : 1). Cette multitude d'effets de sens et de valeurs modales des verbes a intéressé de nombreuses chercheuses, comme le rappelle Le Querler (2004) dans son état de l'art sur la notion de modalité ou encore

¹¹ Pour une réflexion plus détaillée autour de l'interférence entre *modalité* et *mode*, se référer à Monte (2011).

¹² Pour une réflexion plus détaillée sur le temps verbal en tant que facteur de modalité, se référer à Le Querler (2004). Nous reviendrons également sur les valeurs du présent de l'indicatif et du conditionnel présent dans l'introduction du chapitre 7 de ce travail.

¹³ Nous reviendrons sur la terminologie des notions de *valeur modale* et d'*effet de sens* dans la partie 4.1. de ce travail.

Flaux et Lagae (2014) dans leur préface du numéro de la revue *Langages* intitulé *Syntaxe et sémantique des marqueurs modaux*.

C'est ce caractère polysémique des verbes modaux qui fait l'objet de nombreuses analyses dans les recherches sur la modalité¹⁴. Le plus souvent, ces études concernent les valeurs modales que ces verbes peuvent exprimer. C'est pour cela qu'on ne peut aborder les verbes modaux sans mentionner les catégories modales qu'ils expriment.

Ces deux notions sont intrinsèquement liées : les verbes modaux expriment des modalités, et sont donc définies par elles, et ces dernières se trouvent transcrites à travers les verbes modaux.

Nous pouvons résumer cette interrelation avec Chu (2008) qui explique que

[l']extrême richesse des modalités qu'on trouve à travers les emplois des verbes modaux est en fait le résultat d'une interférence entre le cadre de la modalisation [...] et le sémantisme de chaque verbe modal. D'un côté, le sémantisme des verbes modaux les destine à certaines positions dans le cadre de la modalisation : de l'autre, le sens qu'on obtient pour un verbe modal dans un énoncé provient d'une prise de son sémantisme dans les catégories du cadre de la modalisation. (Chu 2008 : 86)

Cette interrelation entre les verbes modaux et les modalités est digne du paradoxe de l'œuf et de la poule. Notre thèse n'a pas pour objectif de résoudre cette question, mais elle offre une entrée pertinente pour aborder les problématiques liées à la sémantique des verbes modaux : si les notions de modalités dépendent des verbes modaux, et vice-versa, comment déterminer et/ou étudier les traits sémantiques caractéristiques de ces verbes ?

Certaines autrices évitent d'aborder la sémantique lorsqu'elles définissent ce qu'est ou non un verbe modal. Elles se concentrent davantage sur leur nature syntaxique particulière. Huot (1974), par exemple, définit les verbes modaux « en français contemporain, [comme] une série de verbes qui, en structure de surface, peuvent être directement suivis d'un infinitif » (Huot 1974 : 9). Leeman-Bouix (2002) met également au premier plan la structure de la phrase lorsqu'elle classe les verbes modaux dans la catégorie des semi-auxiliaires (Leeman-Bouix 2002 : 118). En intégrant les verbes modaux dans la catégorie des (semi-)auxiliaires, nous pourrions ainsi éviter la question de leur noyau sémantique et des catégories modales car ces types de verbes sont décrits comme « “lexicalement vides” [et] comptés comme [des] “outils grammaticaux” » (Blanche-Benveniste 2002 : 75). De ce point de vue, les verbes modaux seraient donc uniquement des éléments fonctionnels qui permettent à la locutrice de « rendre le

¹⁴ Dans une approche diachronique, Dell'Oro (accepté-a) explique que ce sont également les caractéristiques sémantiques des marqueurs modaux qui sont étudiés, et plus précisément les changements de sens à travers le temps.

sens d'un verbe plus approprié aux circonstances d'une communication [...] » (Chu 2008 : 7). Ce type de définition admet une grande liste de verbes répondant à ces critères¹⁵.

Un autre élément qui plaide en faveur des verbes modaux dans les études sur la modalité est un critère purement quantitatif : les verbes modaux sont très fréquents dans les conversations quotidiennes. Ceci peut également expliquer l'intérêt qu'ils suscitent dans les recherches sur la modalité :

La communication linguistique n'est pas un simple échange d'informations et bien des facteurs interpersonnels entrent en jeu dans le dialogue : on doute, on émet une réserve, on suggère, on pronostique, on cherche à partager une croyance, on cède, on laisse voir un événement sous un aspect particulier... Les verbes modaux sont impliqués directement dans cette dynamique communicative et leurs emplois fréquents et variés dans les communications verbales donnent la mesure de la complexité de nos échanges langagiers. (Chu 2008 : 153)

Malgré cette importance et la profusion des études à leur propos, les verbes modaux ont encore rarement fait l'objet d'études quantitatives. En effet, l'émergence de la linguistique de corpus dans les années 1950¹⁶ n'a pas profité aux recherches sur les caractéristiques sémantiques des verbes modaux. Pourtant, dès les premières publications faisant référence aux méthodologies quantitatives, il a été relevé qu'elles pourraient être très utiles pour l'analyse de la sémantique (Oakey 2022 : 185). Cela est d'autant plus intéressant lorsqu'on applique le précepte très simple et néanmoins d'une grande importance de Firth qui explique que « You shall know a word by the company it keeps ! » (Firth 1957 : 11)¹⁷. Et comment mieux appréhender l'environnement d'un mot que par des outils statistiques qui nous révèlent le co(n)texte¹⁸ privilégié du mot en question ?

Dans cette thèse, nous emploierons ainsi les outils de la linguistique de corpus et de la textométrie afin d'appréhender au mieux les propriétés sémantiques des verbes modaux. Plus précisément, nous emploierons les logiciels TXM (Heiden et al. 2010) et Hyperbase (Brunet 2011) pour étudier les préférences co(n)textuelles de ces verbes dans deux corpus : un corpus d'articles de presse tirés du quotidien français *Le Monde* durant l'année 2010 ainsi qu'un corpus encyclopédique dont les textes proviennent de l'encyclopédie participative Wikipédia. Ces préférences d'apparition seront calculées et analysées avec les cooccurrences significatives que les verbes modaux présentent avec les sujets grammaticaux qui les régissent ainsi que les verbes

¹⁵ Chu (2008) en considère d'ailleurs 12 dans son ouvrage sur les verbes modaux du français. Nous reviendrons sur cette liste plus en détail dans le chapitre 2 de ce travail.

¹⁶ Nous reviendrons sur cet aspect dans la partie 6.1. de ce travail.

¹⁷ Au vu de la popularité de cette formulation, nous avons décidé de ne pas la traduire contrairement aux autres citations qui figurent dans cette thèse.

¹⁸ Nous reviendrons sur la différence entre cotexte et contexte dans la partie 4.1. de ce travail.

à l'infinitif qu'ils modalisent. Trois méthodes de calcul seront utilisées pour mesurer ces cooccurrences, soit l'indice de spécificité, l'analyse factorielle des correspondances ainsi que l'analyse arborée¹⁹.

Dans les pages suivantes, nous examinerons comment les chercheuses ont abordé la question des caractéristiques sémantiques des verbes modaux. Cette revue nous permettra de présenter nos propres critères de sélection pour le paradigme des verbes modaux choisis dans cette thèse, soit *pouvoir*, *devoir*, *falloir* et *vouloir*. Ensuite, nous détaillerons l'organisation de notre recherche et la structure de cette thèse.

¹⁹ La méthodologie sera explicitée en détail dans le chapitre 6 de cette thèse.

2. Quels verbes modaux considérer ?

Opter pour un cadre théorique délimité (notamment les catégories modales que nous considérerons ou encore la méthodologie avec laquelle nous traiterons les verbes modaux) est nécessaire car définir une classe fermée de verbes modaux est un exercice laborieux :

Contrairement à ce qui pourrait apparaître au premier abord, établir une classe de verbes modaux n'est pas chose facile. Les ouvrages de linguistique et les grammaires du français parlent souvent des verbes modaux ou des auxiliaires modaux, mais on n'est jamais arrivé à un consensus sur la délimitation d'une telle classe. Le désaccord sur les membres de cette classe est aussi général que celui sur la façon d'appeler cette classe. (Chu 2008 : 22)

Le problème de délimitation vient des critères à prendre en compte : sont-ils d'ordre sémantique, morphologique ou syntaxique ? Selon le domaine linguistique dans lequel se situent les chercheuses – mais aussi selon la vision de la modalité qu'elles adoptent – les verbes modaux analysés ne sont pas les mêmes. Dans notre cas, nous avons fait le choix d'opter pour des critères sémantiques.

Comme énoncé précédemment, les verbes modaux les plus prototypiques sont *pouvoir* et *devoir*. Ils peuvent être étudiés tant ensemble, et donc comparés (Barbet 2012, 2015 : De Saussure 2014a : Sueur 1977, 1979 : Vetters 2004, 2012), qu'analysés individuellement, en mettant en parallèle les différents effets de sens qu'ils peuvent revêtir (pour *pouvoir* : Tasmowski et Dendale (1994) ; pour *devoir* : Dendale (1994, 2000) ; Huot (1974) : Kronning (1990, 1996, 2001) ; Rossari et al. (2007) ; Rossari et al. (2018a) ; Vetters et Barbet (2006)). Comme nous le disions plus haut, cette profusion d'études à propos de *pouvoir* et *devoir* est due au fait que ce sont les deux verbes prototypiques d'une définition étroite de la modalité « héritées des traditions aristotéliennes et centrées sur les notions de *nécessaire* et de *possible* » (Barbet 2013 : 4).

D'autres verbes peuvent être listés comme modaux. C'est le cas lorsqu'une conception large de la modalité est acceptée, conception qui prend en compte toute trace de la locutrice dans le paradigme des marqueurs modaux (Le Querler 2004 : 646). Dans cette conception large, la locutrice peut donc aussi manifester « [...] son doute, son appréciation ou bien sa volonté à propos d'un contenu propositionnel » (Le Querler 1996 : 14) et ce, avec des expressions aussi variées qu'il existe de catégories grammaticales en français.

Dans la figure 1, nous avons regroupé plusieurs recherches qui analysent plus de deux verbes modaux²⁰ (Gaätone in Blanche-Benveniste 2002 ; Caron & Caron-Pargue 2003 ; Chu 2008 ; Fontanille in Colas-Blaise 2022 ; Flaux & Lagae 2014 ; Sandfeld et Ruwet in Huot 1974 ; Leeman-Bouix 2002 ; Riegel et al. 2008). La liste la plus longue est celle de Gaätone (in Blanche-Benveniste 2002) qui comporte 17 verbes. Sur les neuf analyses que nous citons ici, nous pouvons établir une moyenne de huit verbes modaux par travail.

²⁰ Nous voulions ici montrer quels sont les verbes modaux considérés par les linguistes ayant travaillé sur les verbes modaux du français. En traitant des listes comportant minimum trois verbes, nous pouvons faire ressortir les critères de sélection et donc de définition de ce qu'est un verbe modal ou non. Les recherches qui se focalisent uniquement sur un ou deux verbes sont plus ciblées sur les caractéristiques individuelles des verbes traités et non pas sur ce qui les rassemblent ou différencient.

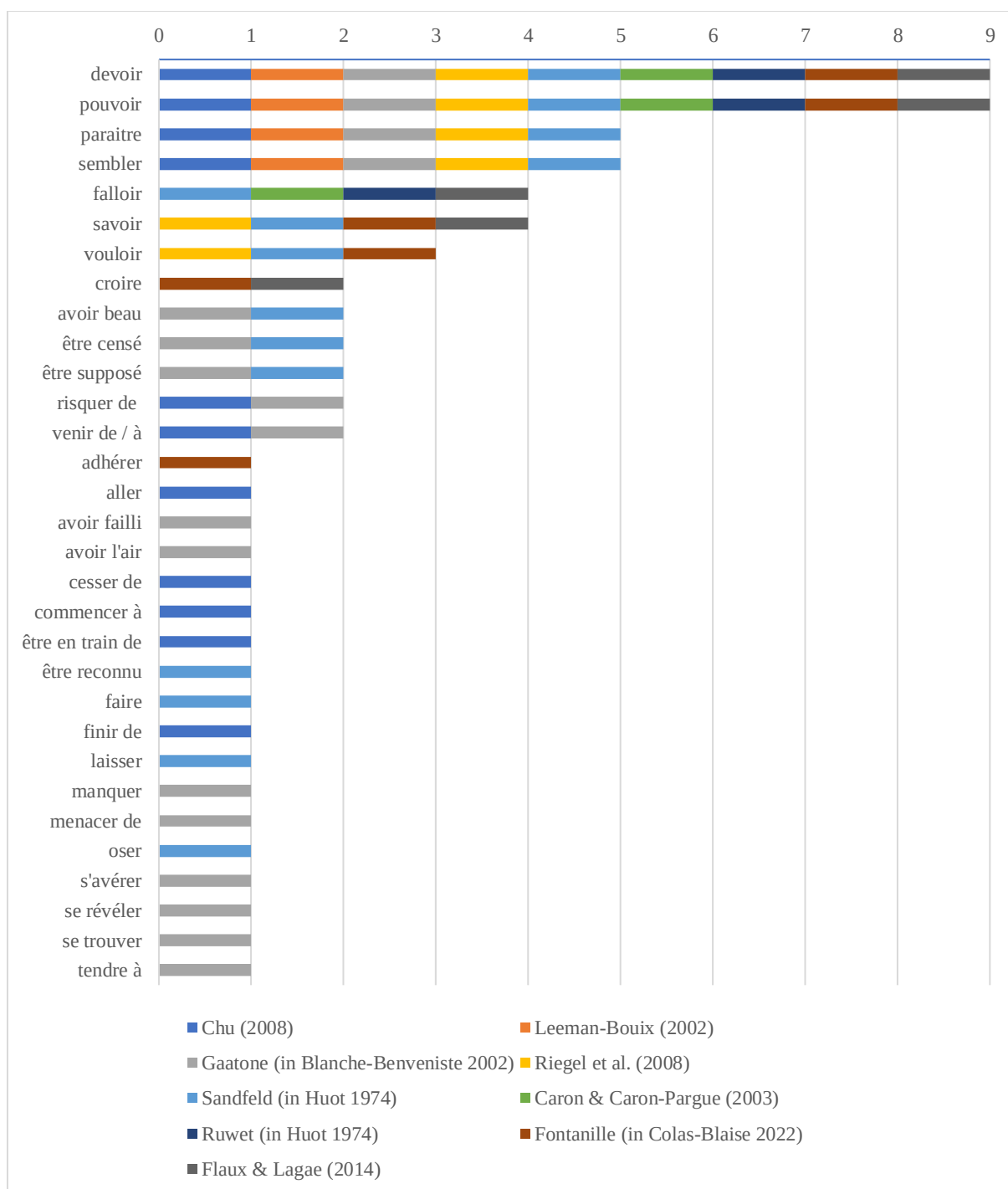


Figure 1 : Les verbes modaux considérés dans les recherches qui traitent plus de deux verbes modaux.

Sans surprise, ce sont *pouvoir* et *devoir* qui sont le plus souvent analysés dans des recherches sur les verbes modaux. Toutes les autrices de la figure 1 les intègrent à leurs études. Suivent *sembler* et *paraître* dans cinq listes, puis *falloir* et *savoir* dans quatre, et *vouloir* dans trois. Ainsi, les autrices semblent la plupart du temps se focaliser sur des verbes pouvant traduire une modalité épistémique reflétant « la vérité dite *subjective* – selon le point de vue souvent adopté en linguistique – ou le savoir – selon le point de vue logique. Les modalités épistémiques

marquent notamment le domaine du savoir, de la croyance, du doute » (Barbet 2013 : 14). Exception toutefois pour *devoir*, aussi pris en compte par de nombreuses autrices et traditionnellement défini comme traduisant la *nécessité* et l'*obligation*, à savoir les modalités déontiques.

Chu (2008) se démarque particulièrement par sa longue liste, accordant une grande place aux verbes qui reflètent des caractéristiques sémantiques liées à la temporalité des événements. Ce n'est pas surprenant de prendre en compte le facteur *temps* dans le cadre des études sur les verbes modaux, mais, habituellement, c'est davantage le *temps* en tant que constituant de la conjugaison qui est considéré. Flaux et Lagae (2014) précisent que « [p]our le français, ce sont évidemment les marqueurs flexionnels du futur, du conditionnel et de l'imparfait qui ont le plus retenu l'attention des linguistes [...] » (Flaux & Lagae 2014 : 4) dans le cadre des études sur les marqueurs modaux. L'imparfait et le conditionnel donnent l'occasion à la locutrice d'exprimer son doute, et donc la modalité épistémique (De Saussure 2012 : 132 ; Dendale 1994 : 37), alors que le futur exprime un ordre, donc une modalité déontique. Ce dernier point est très bien démontré par Le Querler (2004) :

Par exemple, le futur simple dans *Vous préparerez ce dossier avec soin* fait bien sûr référence à une époque d'un futur proche où l'action se réalisera, mais établit surtout un rapport intersubjectif marquant la modalité de l'ordre. Le futur ici marque une relation d'autorité entre le locuteur et l'interlocuteur de façon analogue à un impératif, par exemple (« Préparez ce dossier avec soin »). Temporalité et modalité dans cet énoncé sont très liées et le tiroir verbal du futur exprime donc bien autre chose que la temporalité, ici la modalité intersubjective [déontique]. (Le Querler 2004 : 651)

Il n'est donc pas rare d'avoir des études sur les marqueurs modaux qui prennent en compte les notions de *temps* ou de *temporalité*. Selon certaines autrices, ces notions sont également une trace de la subjectivité de la locutrice :

La temporalité est une notion construite autour d'un moment qui est le *maintenant* du locuteur²¹. Le moment de l'action ou de l'état de choses est repéré par rapport au moment de renonciation. Le moment d'énonciation est, en fait, un moment fictif. En effet, c'est le point 0 d'un axe temporel : avant, c'est le passé ; après, c'est le futur. Mais le point zéro est insaisissable et variable en extension : le présent peut aussi bien être constitué d'une seconde, d'une heure, d'une journée, d'une année, d'une période quelconque... (Le Querler 2004 : 650)

Dans notre étude, nous n'élargirons pas notre définition de la modalité aux marqueurs temporels. Cependant, nous estimons qu'il est pertinent de considérer d'autres verbes que *pouvoir* et *devoir* dans notre paradigme. En effet, *pouvoir* et *devoir* sont les verbes

²¹ Nous mettons en évidence.

prototypiques des études sur la modalité car ils sont les exemples les plus marquants des modalités épistémiques et déontiques. Comme susmentionné, ils sont aussi souvent comparés puisqu'ils véhiculent ces deux modalités avec des nuances sémantiques plus ou moins marquées²².

Toutefois, ces caractéristiques ne sont pas propres à ces verbes modaux. D'autres verbes de la figure 1 ont également la particularité de pouvoir exprimer tant une modalité déontique (4a-d) qu'épistémique (5a-d) :

4a. Tu **dois** ranger ta chambre.

4b. **Peux**-tu ranger ta chambre ?

4c. Il **faut** ranger ta chambre.

4d. Je **veux** que tu ranges ta chambre.

5a. Il **doit** être bête pour croire que la terre est plate.

5b. Il se **peut** qu'il soit bête pour croire que la terre est plate.

5c. Il **faut** être bête pour croire que la terre est plate.

5d. Il **veut** pleuvoir ce soir.

5d est quelque peu particulier car il montre un usage épistémique du verbe *vouloir* attesté seulement dans quelques régions francophones et uniquement dans certains contextes²³.

D'autres verbes souvent considérés comme modaux ne peuvent pas se décliner dans ces deux valeurs modales. C'est le cas notamment de *savoir* qui peut exprimer une *capacité*, donc plutôt une modalité descriptive mais non pas une modalité déontique. *Croire*²⁴, *sembler* et *paraître* traduisent généralement la modalité épistémique mais ne peuvent pas véhiculer des valeurs déontiques.

Dans notre paradigme, les verbes modaux pris en considération présentent certes des similarités sémantiques, mais aussi des différences. Le fait que *pouvoir*, *devoir*, *falloir* et *vouloir* peuvent traduire les deux modalités les plus traditionnelles et qu'ils sont plus ou moins interchangeables dans certains contextes nous pousse à les examiner davantage, à étudier ainsi leurs similarités mais surtout leurs différences sur le plan sémantique.

²² Nous verrons dans la partie 4.6.1. que c'est surtout sur le plan épistémique qu'ils sont comparés et que *devoir* véhiculerait ainsi une certitude plus importante que *pouvoir* dans des co(n)textes similaires.

²³ Gadet (2021) parle plutôt d'un *vouloir* de futur périphrastique dans ce type de construction. Nous détaillerons cet emploi dans la section 4.5.1. et expliquerons ainsi pourquoi nous pouvons considérer cet usage comme un exemple de modalité épistémique.

²⁴ De plus, *croire* est plus souvent considéré comme un verbe de *pensée* (Vet 1994), d'opinion (Martin 1988) ou encore d'assertion (Borillo 1982).

3. Articulation de la recherche

À partir de ces premières réflexions théoriques, nous pouvons dresser l'articulation de la recherche qui sera menée dans cette thèse.

Étant donné que le point de départ de notre catégorisation des verbes modaux se fonde sur les caractéristiques sémantiques de ceux-ci, nous allons, dans un premier temps, proposer une réflexion plus approfondie sur la notion de **polysémie** des verbes modaux et sur ses implications dans notre étude (4.1.). Nous ferons également le point sur la notion de grammaticalisation parce qu'elle concerne tous les verbes modaux que nous étudions ici, et qu'une « perte de sens » est finalement une caractéristique sémantique (4.1.1). Ces considérations nous permettront de dresser les premiers points communs ou les éventuelles divergences que présentent *pouvoir*, *devoir*, *falloir* et *vouloir* concernant l'expression de leurs effets de sens, leurs valeurs modales ou encore leurs noyaux sémantiques respectifs. Rappelons que, depuis l'analyse sémique de Pottier (1964) et les réflexions sur le langage de Wittgenstein (1961), nous savons que c'est par comparaison entre unités lexicales aux noyaux sémantiques et effets de sens proches que nous pouvons faire ressortir les spécificités de chacune. Il nous semble ainsi intéressant de prendre en compte ces quatre verbes qui partagent des propriétés communes pour faire ressortir leur unicité.

Nous nous attarderons ensuite sur chacun de ces verbes modaux et brosserons leur portrait sémantique à partir de la littérature existante à leur sujet. Nous mettrons en évidence leurs **noyaux sémantiques** (4.2. pour *pouvoir*, 4.3. pour *devoir*, 4.4. pour *falloir* et 4.5. pour *vouloir*), leurs divers **effets de sens** (4.2.2. pour *pouvoir*, 4.3.3. pour *devoir*, 4.4.3. pour *falloir* et 4.5.4. pour *vouloir*) ainsi que les **valeurs modales** (4.2.3. pour *pouvoir*, 4.3.2. pour *devoir*, 4.4.4. pour *falloir* et 4.5.4. pour *vouloir*) qu'ils peuvent véhiculer. Pour certains de ces verbes, ces différents paramètres sont déterminés à l'aide du co(n)texte, et plus particulièrement le type de **sujet grammatical** qui les régit ou encore **la complémentation verbale sous forme de verbe à l'infinitif** qui suit le modal. Nous allons ainsi relater les influences de ces paramètres sur la sémantique des verbes modaux (4.2.4. pour *pouvoir* et 4.3.6. à 4.3.8.2. pour *devoir*). Nous présenterons aussi les différentes méthodologies qui ont été employées pour faire ressortir les caractéristiques sémantiques de ces verbes modaux (4.3.9. pour *devoir*, 4.4.6. pour *falloir* et 4.5.7. pour *vouloir*). Nous finirons notre tour d'horizon des caractéristiques sémantiques des verbes modaux sur les études comparatives qui existent à leur sujet (4.6.) et, à la suite de cela, nous exposerons nos questions et hypothèses de recherches (5).

Ces réponses seront amenées à l'aide des outils tirés de **la linguistique de corpus** (6.1.). Il sera notamment question des notions de **corpus** (6.3.1.), de **fréquence** (6.3.2.), de **cooccurrence** (6.3.3.), ou encore d'**empan** (6.3.4.). La façon de présenter les résultats sera également mise en avant, notamment à l'aide de l'**indice de spécificité** (6.4.1.), l'**analyse factorielle des correspondances** (6.4.2.) ou encore de l'**analyse arborée** (6.4.3.). Toujours dans le cadre du chapitre 6 consacrée à la méthodologie, nous présenterons ensuite nos corpus et les divers logiciels que nous utilisons pour mener à bien nos recherches (6.5.). Il sera ainsi question de **TXM** (6.5.2.1.) et **Hyperbase** (6.5.2.2.).

À la suite de cela, nous présenterons les résultats de nos analyses. Nous nous attarderons dans un premier temps sur les combinaisons des verbes modaux avec les sujets grammaticaux sous forme de noms (7.1.). La seconde partie de notre analyse concerne la combinaison des modaux avec la complémentation verbale sous forme de verbe à l'infinitif (7.2.). Ces recherches permettent ainsi de mettre en exergue les propriétés déjà attribuées à ces verbes modaux, mais aussi de les remettre en cause, notamment le fait que ces éléments de la valence exercent une influence sur l'interprétation sémantique du modal. Ces résultats nous révéleront également de nouvelles caractéristiques sur ces quatre verbes modaux. Nous résumerons notre recherche et les perceptives que celle-ci amène à la fin de la thèse (8).

4. La sémantique des verbes modaux

A cause de leurs valences particulières, les verbes modaux sont parfois considérés comme sémantiquement vides de sens : « On les définit parfois comme des verbes qui, sans être des auxiliaires suivis de participes passés, ont cette propriété de s'accorder en nombre et en personne avec un apparent sujet mais de n'exercer aucune sélection sur les valences » (Blanche-Benveniste 2002 : 84). Cette remarque rejoint la classification d'Abeillé et Koenig (2021a : 169) qui ont catégorisé les verbes demandant un complément sous forme de verbe à l'infinitif dans deux catégories :

Les linguistes distinguent deux types de verbes avec infinitif : ceux comme *vouloir*, souvent dits à *contrôle*, et ceux comme *sembler*, souvent dits à *montée*. Nous les appelons respectivement verbes à *partage* et verbes à *héritage* : dans le premier cas, le sujet de l'infinitif est sélectionné à la fois par le verbe principal et l'infinitif, qui le *partagent* : dans le second cas, il est sélectionné uniquement par l'infinitif et *hérité* par le verbe principal. (Abeillé & Koenig 2021a : 169)

Les sujets des verbes de ces deux catégories n'auront ainsi pas la même valeur :

Ainsi, avec un verbe comme *vouloir*, le sujet [...] est interprété comme un agent (celui qui veut). Ce n'est pas le cas avec un verbe comme *sembler* : [dans *Jean semble dormir*,] Jean n'est pas interprété comme 'celui qui semble', mais simplement comme celui qui dort. Du point de vue sémantique, l'argument de *sembler* est l'ensemble sujet + infinitif, c'est-à-dire 'Jean dort'. (Abeillé & Koenig 2021a : 169)

Abeillé et Koenig (2021a) classent *pouvoir* et *devoir* aux côtés de *sembler* dans les verbes à *héritage*, « qui n'assignent pas de rôle sémantique à leur sujet » (Abeillé & Koenig 2021a : 169). *Vouloir*, quant à lui, serait ainsi le seul verbe de notre recherche à être considéré comme un verbe à *partage*²⁵.

Malgré cette distinction de comportement, nous ne pouvons pas faire abstraction du fait que tous les verbes modaux communiquent un sens, ou du moins, un concept lorsqu'ils sont utilisés : « Il est difficile de contester que les verbes modaux encodent des concepts facilement compréhensibles, et certainement très fondamentaux dans la cognition : la possibilité et la nécessité » (De Saussure 2017 : 282). De ce fait, il semble dérisoire de les considérer uniquement comme des auxiliaires et ainsi vides de caractéristiques sémantiques malgré leur

²⁵ Nous nuancerons ces propos dans la section 4.5.1. consacrée à *vouloir* et la question de son intégration dans la catégorie des verbes modaux.

valence particulière. Ce n'est donc pas anodin que certaines chercheuses leur concèdent un noyau sémantique²⁶.

De plus, comme nous l'énoncions précédemment, c'est principalement l'aspect sémantique des verbes modaux qui est analysé dans la littérature à leur propos et non pas leurs particularités syntaxiques. Comme le rappelle Barbet, « [l]a question souvent posée au début des monographies consacrées aux verbes modaux [...] concerne la représentation de leur multiplicité de sens » (Barbet 2013 : 95). Cette thèse ne déroge pas à la règle.

Ainsi, nous approcherons la sémantique des verbes modaux par les notions de *noyaux sémantiques*, *effets de sens* et *valeurs modales* (4.1). Ces trois notions seront ensuite exploitées pour chacun des verbes modaux que nous étudions dans cette thèse, soit *pouvoir* (4.2.), *devoir* (4.3.), *falloir* (4.4.) et *vouloir* (4.5.). Chacun de ces verbes sera traité avec les particularités qui le définissent : de cette manière, *pouvoir* est le seul verbe de notre étude bénéficiant d'une partie sur son emploi post-modal (4.2.5.), *devoir* est le verbe dont la sémantique a été le plus précisée à l'aide du co(n)texte (4.3.3.5 – 4.3.3.8), *falloir* a souvent été étudié d'un point de vue étymologique (4.4.1.) et comparatif entre les langues (4.4.6.1.) et *vouloir* n'est pas toujours qualifié de modal et a donc été étudié aux côtés d'autres verbes de *volition*²⁷ (4.5.1. et 4.5.7.). Il sera également question des phénomènes de grammaticalisation pour chacun de ces modaux (4.3.5. pour *devoir*, 4.4.5. pour *falloir* et 4.5.6. pour *vouloir*). Nous nous focaliserons aussi sur les méthodologies employées jusqu'alors dans les études de ces verbes (4.3.9 pour *devoir*, 4.4.6. pour *falloir* et 4.5.7. pour *vouloir*). Enfin, nous terminerons ce tour d'horizon avec les études comparatives entre ces verbes (4.6.).

4.1. La polysémie des verbes modaux : pas qu'une question de sémantique...

La question de la polysémie des verbes modaux ne date pas d'hier. Ainsi, « [d]ès les premiers travaux sur les verbes modaux en français (cf. Gougenheim 1929 ; Sten 1954 ; Huot 1974 ; Sueur 1975), on a [...] tenté d'organiser cette multiplicité de sens » (Barbet 2013 : 20)²⁸. Pour

²⁶ Nous comprenons ici le terme *noyau sémantique* dans le sens de Pottier (1964), c'est-à-dire comme un ensemble de sèmes centraux auxquels peuvent se greffer d'autres sèmes afin d'en apporter des nuances.

²⁷ La *volition* est un terme plus technique, utilisé notamment en psychologie pour qualifier le processus ou les actes liés à la volonté. Dans cette thèse, nous utiliserons de manière interchangeable le terme *volonté* et *volition* pour traiter les caractéristiques sémantiques de *vouloir*.

²⁸ Dans cette thèse, nous ne reviendrons pas sur les différences entre les approches *homonymiques*, *polysémiques* ou encore *monosémiques* des verbes modaux. Barbet (2013) résume ces distinctions dans le chapitre 4 de sa thèse.

traiter ces aspects de polysémie des verbes modaux, nous pouvons admettre qu'il y a trois notions principales qui nous semblent indispensables à définir :

- I) Il y a tout d'abord le *noyau sémantique*²⁹ (désormais NS) qui se rapporte au sens fondamental que le verbe modal possède et qui peut être retrouvé peu importe l'usage du verbe en question. Ainsi, par exemple, *pouvoir* est généralement défini par un NS de *possibilité* et *devoir* de *nécessité*³⁰.
- II) Ce NS peut ensuite découler sur plusieurs *effets de sens*³¹ qui se traduisent par les notions d'*obligation*, de *possibilité*, de *probabilité*, de *volonté*, etc. Ces effets de sens sont les concepts co(n)textuels véhiculés par les verbes modaux.
- III) Quant à la *valeur modale*³², elle fait référence aux catégories classiques des modalités telles que les modalités *épistémiques*, *déontiques*, *descriptives* ou encore *bouliques* pour ne citer que celles-ci.

Ces notions se regroupent la plupart du temps. Par exemple, la *possibilité* est définie tant comme étant un NS qu'un effet de sens de *pouvoir*, qui est également caractérisée par la valeur modale *épistémique*. Dans ce travail, nous utiliserons la notion de *valeur modale* uniquement lorsque nous mentionnons les catégories modales citées en III, soit les modalités *épistémiques*, *déontiques*, *descriptives* ou encore *bouliques*. Les *effets de sens* dénotent les concepts véhiculés dans ces différentes catégories modales comme la *probabilité*, l'*obligation* ou encore le *désir* pour ne citer que ceux-ci.

Même si la grande majorité des linguistes intéressées par le sujet admettent que les verbes modaux sont polysémiques, les autrices ne s'accordent pas toutes sur la façon dont les effets de sens et/ou valeurs modales varient pour un même verbe. Ainsi, Barbet (2013) et De Saussure (2014a, 2014b) s'accordent à dire que les effets de sens des verbes modaux dépendent du contexte, et qu'ils sont « dérivés pragmatiquement » (Barbet 2013 : 106). Ces autrices admettent un NS pour les verbes *pouvoir* et *devoir* mais c'est bien le contexte, du point de vue pragmatique, qui prévaut sur l'effet de sens exprimé par le verbe modal. D'autres, comme Gosselin (2010), estiment que ces verbes contraignent « sévèrement [leurs] effets de sens

²⁹ Terminologie retrouvée chez Vetters et Barbet (2006), Vetters (2012) ainsi que Goldschmitt et Landvogt (2008). De Saussure (2014a, 2014b), Barbet (2013), Gosselin (2010) et Chu (2008) l'appellent *noyau de sens*.

³⁰ Nous préciserons et questionnerons ces attributions dans les parties consacrées à ces verbes, soit la 4.2.1. pour *pouvoir* et 4.3.1. pour *devoir*.

³¹ Gosselin (2010) et Le Querler (1996) utilisent cette terminologie dans leurs ouvrages sur la modalité.

³² Terminologie employée chez Flaux et Lagae (2014), Riegel et al. (2008), Le Querler (2004), Monte (2011) et Douay (2003).

possibles (sans que ces contraintes soient directement déductibles du noyau de sens) » (Gosselin 2010 : 454). Il y aurait ainsi des effets de sens tellement éloignés du NS que le lien entre ces effets et le NS serait à peine tangible. Quant à Rossari (2020), elle admet que les formes modales possèdent une seule contribution sémantique – qui équivaut au *noyau sémantique* défini ci-dessus – mais que celle-ci peut se décliner en différentes valeurs selon le contexte. Le contexte, dans sa théorie, est à comprendre dans le sens textuel et non pas pragmatique. Elle propose ainsi un modèle d'analyse pour les formes modales qu'elle nomme *triadique* car il est composé de trois paramètres :

- I) Le premier (P1) est la contribution sémantique de la forme (CS) (Rossari 2020 : 19), qui correspond à ce nous appelons ici *noyau sémantique*. Il s'agit, dans notre thèse, d'un NS de *possibilité, nécessité/contrainte* ou encore d'*intention/volition* selon le verbe étudié, comme nous le verrons dans les pages qui vont suivre.
- II) Le deuxième (P2) est le « mode d'application de la CS » (Rossari 2020 : 19). Ce mode peut concerner soit le contenu propositionnel de l'énoncé ou le simple fait qu'il soit énoncé, donc le phénomène même d'*énonciation*. En d'autres termes, il porte soit sur le *dit*, soit sur le *dire*, selon la différenciation opérée notamment par Ducrot (1984). Cela rejoint également la distinction de Nølke (1993) dans son étude sur les traces énonciatives, rappelée par Büyükgüzel (2011) :

Nølke (1993 : 85) propose quant à lui la définition suivante : « Par modalités d'énonciation, j'entends les éléments linguistiques qui portent sur le dire, pour reprendre une expression chère à beaucoup de linguistes. Ce sont les *regards* que le locuteur jette sur son activité énonciative ». Insistant sur la distinction entre les modalités d'énonciation et les modalités d'énoncé, il ajoute que « si les modalités d'énonciation portent sur le dire, les modalités d'énoncé portent sur le dit » (Nølke, 1993 : 143). (Büyükgüzel 2011 : 134)

- III) Enfin, le troisième paramètre (P3) concerne uniquement les cas où le mode d'application porte sur le *dire*. Elle concerne la prise en charge de l'énoncé (PR) qui possède une motivation rhétorique : soit PR est revendiquée, soit elle est mise à distance (Rossari 2020 : 19). Ce paramètre concerne avant tout les énoncés véhiculant une valeur modale épistémique.

Cette distinction entre le *dire* et le *dit*, donc sur une modalité qui porte sur l'*énonciation* ou l'*énoncé*, est commune dans les études sur les marqueurs modaux. L'énonciation, en tant que telle, peut se définir comme

l'acte individuel de production, dans un contexte déterminé, ayant pour résultat un *énoncé* : les deux termes s'opposent comme la *fabrication* s'oppose à l'objet *fabriqué*. L'énonciation est l'acte

individuel d'utilisation de la langue, alors que l'énoncé est le résultat de cet acte, c'est l'acte de création du sujet parlant devenu alors *ego* ou *sujet d'énonciation*. Il s'agit essentiellement, pour les initiateurs de ce concept [...] de dégager les éléments qui, dans les énoncés, peuvent être considérés comme les traces ou les empreintes des *procès d'énonciation* qui les ont produits, puis de dégager leur fonctionnement, leur organisation, leur interaction. (Dubois et al. 2012 : 180-181)

Les verbes modaux font indéniablement partie de ces *traces de l'énonciation* (Kerbrat-Orecchioni 1980 : 32) et c'est pourquoi ils sont souvent considérés dans les études prenant en compte la notion d'énonciation. Dans ce cadre, les autrices différencient l'influence du modal selon sa portée, et cette distinction permet d'interpréter l'effet de sens ou la valeur modale du verbe. Par exemple, pour les formes possédant un NS de *nécessité* – comme *devoir* ou encore *falloir* –, nous pouvons admettre que, lorsque la modalité est appliquée sur le *dit* (sur l'énoncé), le verbe véhicule une valeur modale déontique (5) mais lorsqu'elle est appliquée sur le *dire* (sur l'énonciation) nous aurons davantage une lecture épistémique (6) :

6. On **doit** avoir 18 ans pour voter.

7. Sa voiture est dehors, il **doit** être arrivé.

Nous n'opérons pas une distinction de la portée du verbe modal³³ pour en analyser et lister les effets de sens et valeurs modales mais comme Rossari (2020), nous postulons que les effets de sens et/ou les valeurs modales des formes modales varient selon le contexte. Nous considérons ici le contexte, non seulement au sens pragmatique, mais aussi comme un *cotexte* qui désigne « l'ensemble des “mots” qui précèdent et suivent l'item considéré » (Portine 2017 : 30), définition très adéquate avec la méthodologie de la linguistique de corpus que nous utilisons dans cette thèse. C'est en partie l'importance de l'usage des mots dans un co(n)texte défini qui a contribué au développement de cette méthodologie. Rappelons que Firth – contextualiste qui a mis en avant l'importance des cooccurrences pour étudier le sens des mots et qui a ainsi contribué au développement de la linguistique de corpus³⁴ – avait d'ailleurs écrit : « You shall know a word by the company it keeps » (Firth 1957 : 11).

Cette vision de l'importance du cotexte dans la détermination de l'effet de sens et/ou de la valeur modale du verbe semble rejoindre la perception de quelques chercheuses qui lient explicitement les questions de sémantique des verbes modaux aux notions de syntaxe ou encore

³³ D'ailleurs, avec une méthodologie de linguistique de corpus qui est la nôtre, il peut être difficile et chronophage – mais non pas impossible – d'intégrer ce paramètre dans notre étude.

³⁴ Nous reviendrons plus en détails sur le développement de la linguistique de corpus et l'importance de l'usage dans la partie 6.1. de ce travail.

de valence. Nous pouvons, par exemple, lire chez Domínguez (2004) que « s'il existe un lien intrinsèque entre le sémantisme d'un verbe et son apparition en discours, cette apparition est également le résultat de son fonctionnement syntaxique interne et de son entourage syntaxique externe » (Domínguez 2004 : 897).

Cependant, comme nous verrons dans les nombreux exemples que nous analyserons (chapitre 7), cette vision *cotextuelle* n'est pas toujours suffisante pour déduire l'effet de sens et/ou la valeur modale du verbe. Comme l'a rappelé Recanati (2004), « la sémantique (l'étude formelle de la signification [...]) et la pragmatique (l'étude de la langue en usage) sont désormais conçues comme des disciplines complémentaires, éclairant différents aspects du langage » (Recanati 2004 : 2). Ainsi, même si nous nous insérons dans un cadre qui étudie *a priori* les caractéristiques *sémantiques* des verbes modaux, il est nécessaire de tenir compte des facteurs *pragmatiques* qui peuvent influencer sur l'interprétation de ces propriétés. C'est pourquoi nous devons également aborder la notion d'enrichissement pragmatique qui « prend pour point de départ un élément de ce qui a été dit dans la phrase » (Zufferey & Moeschler 2012 : 114). Recanati (2004) explique cela par les processus d'élaboration ou d'extension du sens :

Comme les mots sont appliqués, dans un contexte donné, à des situations spécifiques, leur sens est ajusté. Selon que le sens conventionnel est entièrement ou partiellement schématique pour la situation discutée, l'ajustement prendra l'une des deux formes suivantes : l'élaboration du sens (enrichissement) ou l'extension du sens (relâchement). Dans l'élaboration du sens, le sens porté par les mots devient plus spécifique grâce à l'interaction avec les facteurs contextuels. Dans l'extension du sens, les dimensions de sens qui entrent en conflit avec les spécifications de la cible sont filtrées, mais elles restent en quelque sorte actives et peuvent générer un sentiment de divergence entre le schéma évoqué et le sens construit en appliquant (partiellement) le schéma à la situation en cours. Ce sentiment, tout comme le conflit qui le sous-tend, varie en intensité. Ainsi, il existe un continuum entre les cas ordinaires d'extension du sens que nous ne percevons pas [...] et les cas plus dramatiques de métaphore dont le caractère non littéral ne peut être ignoré. (Recanati 2004 : 77)

Dans notre cas, les effets de sens et valeurs modales que l'on attribue habituellement dans la littérature aux verbes modaux seront considérées comme des phénomènes d'enrichissement du sens, et celles qui s'en éloignent davantage seront envisagées comme des phénomènes d'extension du sens. Dans le second cas, « [d]eux mécanismes sont ainsi à l'œuvre en fonction de la source : le décodage linguistique et les mécanismes inférentiels. » (Zufferey & Moeschler 2012 : 122). L'inférence, au sens de Grice (1979),

utilise comme données d'entrée une série de prémisses à partir desquelles elle fournit en sortie une série de conclusions, qui sont soit directement déductibles à partir de ces prémisses soit simplement

justifiées par elles. En d'autres termes, l'inférence est une forme de raisonnement qui part de prémisses pour arriver à des conclusions. (Zufferey & Moeschler 2012 : 105)

Ces prémisses peuvent être nos connaissances du monde qui ne sont pas toujours explicites dans les énoncés mais qui sont parfois nécessaires pour les comprendre. Les informations déduites par inférences sont nommées *implicatures* chez Grice (1979). Il existe deux types d'implicatures : les conventionnelles et les conversationnelles. Les premières sont fixes et transparaissent directement dans les caractéristiques sémantiques des unités lexicales employées alors que les secondes dépendent du contexte de l'énoncé en question. Ainsi, ces dernières « ont pour propriété d'être déclenchées par des principes généraux liés à la conversation plutôt que d'être attachées à des mots particuliers de l'énoncé » (Zufferey & Moeschler 2012 : 116).

Ainsi, même en nous intégrant dans un cadre porté sur les caractéristiques sémantiques des verbes modaux, il est essentiel de prendre en compte les théories pragmatiques pour en embrasser toute la complexité. Dans une optique de linguistique de corpus, le cotexte tel que nous l'étudierons – en nous basant sur le sujet grammatical ainsi que sur la complémentation verbale – « réveille » un effet de sens ou une valeur modale innée du verbe modal en question. Cependant, ce cotexte n'est parfois pas suffisant pour interpréter les caractéristiques sémantiques du verbe. Nous devons ainsi considérer le contexte dans un sens plus pragmatique ou le rattacher à une valeur d'ordre pragmatique afin de comprendre l'enrichissement ou l'extension de sens que subit le verbe modal dans certains énoncés. Toutefois, lorsqu'aucune de ces interprétations n'est envisageable, nous pouvons également faire appel à la notion de grammaticalisation pour analyser les caractéristiques sémantiques des verbes modaux.

4.1.1. La grammaticalisation des verbes modaux

La grammaticalisation, dans sa définition la plus sommaire, « consiste en l'élargissement de la portée d'un morphème passant d'un statut lexical à un statut grammatical ou d'un statut moins grammatical à un statut plus grammatical [...] » (Narrog & Heine 2011 : 3). Dans le cas des verbes modaux, ce phénomène surgit lorsque le NS du modal n'est pas ou peu récupérable car la locution ou la construction qu'il compose « vide » entièrement ou partiellement le modal de propriétés sémantiques.

Dans son ouvrage, Dostie (2004) rappelle que c'est Meillet qui a défini en premier la grammaticalisation qui est « un phénomène d'évolution, caractérisé par le passage d'une unité lexicale pleine (qu'il appelle « mot autonome ») au statut d'unité grammaticale » (Dostie 2004 :

22). Ce qui est important dans cette citation – mais aussi pour toutes les autrices travaillant sur cette notion – est le fait que la grammaticalisation « est un terme générique qui désigne un processus d'évolution » (Dostie 2004 : 26). En d'autres termes, cette notion « désigne un processus dans lequel quelque chose devient ou est rendu grammatical » (Lehmann 2015 : 11). Le fait que la grammaticalisation soit définie comme un processus montre que, non seulement il y a des étapes, mais, qu'en plus, le statut qu'acquiert une unité lexicale n'est pas figé et peut ainsi encore être amené à évoluer.

Ces étapes de grammaticalisation peuvent être plus ou moins listées. Lehmann (2015) propose les phases suivantes :

Ainsi, nous supposons que la grammaticalisation commence par une collocation libre de mots lexicaux potentiellement non fléchis dans le discours. Celle-ci est convertie en une construction syntaxique par la syntacticisation, par laquelle certains des lexèmes assument des fonctions grammaticales de sorte que la construction peut être qualifiée d'analytique. La morphologisation, qui ici signifie la même chose que l'agglutination, réduit la construction analytique à une construction synthétique, de sorte que les formatifs grammaticaux deviennent des affixes agglutinants. Dans la phase suivante, l'unité du mot est resserrée, car la technique morphologique passe de l'agglutination à la flexion. Cette transition de la morphologie à la morphophonémie sera ici appelée démorphémisation. (Lehmann 2015 : 16)

La grammaticalisation passe donc automatiquement par les différentes unités de la langue, c'est-à-dire du discours à la syntaxe, de la syntaxe à la morphologie pour enfin atteindre un stade phonétique (Lehmann 2015 : 15). Dans le cadre de l'étude de ces phénomènes avec les verbes modaux, nous verrons que nous nous situons principalement entre les phases de discours et de syntaxe.

Heine (2002) propose également quatre étapes de grammaticalisation mais fondées cette fois-ci sur le contexte et les caractéristiques sémantiques :

[...] l'émergence de nouveaux sens grammaticaux peut être décrite au moyen d'un scénario en quatre étapes [...]. À l'étape I, il y a une expression avec un sens « normal » ou un sens « source » apparaissant dans une série de contextes différents. À l'étape II, il y a un contexte de transition donnant lieu à une inférence selon laquelle, plutôt que le sens « source », il y a un autre sens, le sens « cible », offrant une interprétation plus plausible de l'énoncé concerné. À l'étape III, il y a un nouveau type de contexte, le contexte de « basculement », qui ne permet plus une interprétation en termes de sens « source ». Les contextes de « basculement » peuvent être vus comme un dispositif de filtrage qui exclut le sens « source ». Enfin, n'étant plus associé au sens « source », le sens « cible » est désormais ouvert à de nouvelles manipulations : il est libéré des contraintes contextuelles qui l'ont fait émerger, c'est-à-dire qu'il peut maintenant être utilisé dans de nouveaux

contextes. Je me référerai à cette situation comme l'étape IV de conventionnalisation. (Heine 2002 : 85-86)

A partir de l'étape III de Heine (2002), nous pouvons ainsi comprendre cela comme des contextes qui ne font plus apparaître le NS de l'unité lexicale en question, qui peut s'apparenter ici au sens « source ». Nous verrons que concernant les verbes modaux, la plupart des étapes de grammaticalisation ne dépassent pas l'étape II. Heine (2002) précise que ces étapes peuvent soit survenir l'une après l'autre – donc de manière diachronique – mais aussi co-exister – il y aurait ainsi des phénomènes de grammaticalisation synchroniques (Heine 2002 : 86). Les étapes I à III coexistent par exemple pour *vouloir* comme nous le verrons dans les exemples que nous analyserons.

Diewald (2002 : 103-104) propose une analyse similaire mais avec seulement trois étapes – elle omet la phase préliminaire qui apparaît chez Heine (2002) – étapes également liées aux contextes d'utilisation. La première, nommée « diversité d'emplois contextuels », est caractérisée par une expansion non spécifique de la distribution d'une unité lexicale vers des contextes où elle n'avait pas été utilisée auparavant. Ces nouveaux contextes, qualifiés d'« atypiques », permettent l'émergence d'une nouvelle signification en tant qu'implicature conversationnelle – d'où l'importance de considérer les théories pragmatiques pour comprendre les caractéristiques sémantiques des unités lexicales. La deuxième étape marque le déclenchement proprement dit du processus de grammaticalisation. Cette phase est liée à l'apparition d'un type particulier de contexte, appelé « contexte critique », qui se distingue par de multiples ambiguïtés structurelles et sémantiques, ouvrant ainsi la voie à plusieurs interprétations possibles, dont la nouvelle signification grammaticale. Enfin, la troisième étape illustre l'achèvement ou la consolidation du processus de grammaticalisation. Durant cette phase, la nouvelle signification grammaticale se dissocie de l'ancienne signification lexicale. Cette séparation est facilitée par le développement de contextes isolants pour chacune des significations, favorisant une lecture exclusive pour chaque contexte. Une fois cette opposition entre contextes établie, la nouvelle signification grammaticale devient un effet de sens indépendant des autres déjà existants, rendant le processus irréversible et marquant ainsi la polysémie de l'élément grammaticalisé.

Tant la théorie de Lehmann (2002) que celle de Heine (2002) ou de Diewald (2002) démontrent l'importance du contexte pour étudier les caractéristiques sémantiques voire pragmatiques de la grammaticalisation d'une unité lexicale. Le co(n)texte est ainsi primordial que ce soit pour étudier les effets de sens plus « typiques » des verbes modaux que ceux qui sont de l'ordre de la grammaticalisation.

Heine (2002), en plus des étapes susmentionnées, précise également les facteurs propices à la grammaticalisation : le contexte d'utilisation, la fréquence d'utilisation, les processus de raisonnement – en d'autres termes, si une inférence est nécessaire pour comprendre l'unité lexicale dans des contextes précis –, les mécanismes de transfert (métaphore, métonymie...), l'abstraction ou la concrétisation de l'unité lexicale en question, mais aussi, les implications sémantiques, c'est-à-dire, les cas de « perte » de propriétés sémantiques ou de processus de généralisation dans l'utilisation de l'unité lexicale dans un contexte donné (Heine 2002 : 84). Cette « perte » de propriétés sémantiques peut s'apparenter à la notion de « désémantisation » (*bleaching* en anglais), comme l'explique Lehmann (2015) :

La grammaticalisation dépouille les caractéristiques lexicales jusqu'à ce qu'il ne reste plus que les caractéristiques grammaticales. Par conséquent, la fonction relationnelle d'un élément est normalement conservée, tandis que la plupart des caractéristiques sémantiques originales sont perdues lors de la grammaticalisation. (Lehmann 2015 : 138)

Dans le sens que l'entend Lehmann (2015), nous pouvons comprendre que dans certaines situations – ou lors de certaines étapes de la grammaticalisation –, le NS n'est plus perceptible. Cependant, cette « perte » peut être nuancée, ou du moins, plus ou moins conséquente dans certains contextes. Lehmann (2002) donne l'exemple de la préposition à base *de*, dans laquelle nous pouvons encore déceler le NS du nom *base* qui compose cette construction. Ainsi, plus qu'un phénomène parallèle, la désémantisation semble être une caractéristique voire une phase de la grammaticalisation en tant que telle.

Dostie (2004) s'oppose d'ailleurs à ce concept de désémantisation, car « [d']une certaine façon, on pourrait même dire qu'en se grammaticalisant ou en se pragmatiquisant, une unité devient souvent sémantiquement plus complexe » (Dostie 2004 : 39). Ceci est dû, selon elle, au fait que c'est un nouvel effet de sens qui naît parmi ceux déjà existants et qu'on ne peut ainsi pas parler d'appauvrissement pour ce type de phénomène.

Dans cette citation, nous pouvons constater que Dostie (2004) opère une différenciation entre la grammaticalisation et la pragmatiquisation. Elle explique que

[d']une part, une unité lexicale peut développer des emplois grammaticaux : elle aura alors été soumise à un processus de « grammaticalisation ». D'autre part, une unité lexicale / grammaticale peut développer des emplois où elle ne joue pas un rôle sur le plan référentiel, mais bien, sur le plan conversationnel : elle sera alors le résultat d'un processus de « pragmatiquisation ». (Dostie 2004 : 27)

En ce sens, ces unités peuvent ainsi devenir des « signifiants textuels » (Traugott in Lehmann 2015 : 154) et donc être considérés comme des marqueurs discursifs (désormais MD). Les MD sont envisagés comme des marqueurs pragmatiques car

ce sont plutôt les capacités cognitives de l'interlocuteur qui se trouvent sollicitées grâce à eux, ce qui est à rapprocher du constat suivant lequel plusieurs réseaux de MD se sont constitués à partir de verbes de connaissance ou de verbes qui ont développé une telle valeur (verbes de perception, notamment *écouter*, *entendre*, *regarder* et *voir* ; verbes de parole, notamment *admettre*, *dire* et *parler* ; verbes directement cognitifs, notamment *croire*, *comprendre*, *penser*, *savoir* et *sembler*). (Dostie 2004 : 12)

Nous verrons dans nos exemples (chapitre 7) que les verbes modaux que nous analysons se combinent très souvent avec ces types de verbes cités par Dostie (2004), ce qui explique pourquoi il est nécessaire de traiter ici la thématique de la grammaticalisation et de la pragmatification lorsqu'on travaille sur les caractéristiques sémantiques des verbes modaux. Cependant, nous constaterons que nous ne pouvons pas considérer les constructions dans lesquelles les verbes modaux figurent automatiquement comme des MD. En effet, celles-ci doivent répondre à plusieurs critères pour être considérées comme tels. Le premier de ces critères concerne le caractère facultatif des MD sur le plan syntaxique. Ainsi, leur absence n'apporterait pas d'agrammaticalité à l'énoncé (Dostie 2004 : 44). De plus, un MD ne contribuerait pas au contenu propositionnel de l'énoncé (Dostie 2004 : 44), et n'aurait ainsi pas de valeur sémantique. Nous verrons que ce n'est pas automatiquement le cas lorsque les verbes modaux se combinent à des verbes de perception, de parole ou cognitifs comme le suggère la citation susmentionnée. La dernière caractéristique que Dostie (2004) énumère pour considérer une construction comme un MD est d'ordre pragmatique. Ainsi, les MD

peuvent notamment indiquer le début, la continuation ou la fin d'un tour de parole, indiquer une rupture thématique, structurer la conversation et ancrer les énoncés au contexte d'énonciation [...]. Hansen (1998a) insiste également sur le fait que les MD remplissent une fonction métadiscursive et qu'ils fournissent des instructions au coénonciateur sur la façon dont il faut intégrer l'unité hôte du marqueur par rapport à une représentation mentale du discours. (Dostie 2004 : 45)

Nous découvrirons que la fonction métadiscursive peut particulièrement s'appliquer aux exemples que nous analyserons pour les verbes modaux.

Dans cette section, nous avons comme objectif de présenter les caractéristiques principales des phénomènes de grammaticalisation afin de pouvoir opérer une analyse adéquate pour certains usages des verbes modaux, que ce soit dans les parties théoriques qui vont suivre ou dans l'analyse que nous présenterons dans le chapitre 7 de ce travail. Nous ne sommes pas entrée en

détail dans toute la littérature de la grammaticalisation mais les éléments que nous avons relevés nous permettront de soulever des propriétés sémantiques inédites dans le cadre de l'étude des verbes modaux ou, du moins, offrir une nouvelle vision sur des phénomènes déjà relevés mais peu étudiés.

4.2. Pouvoir

Pouvoir est un des verbes les plus fréquents de la langue française (Chu 2008 : 139). Nous nous concentrons dans cette partie sur les caractéristiques sémantiques de ce verbe³⁵, en commençant par son NS (4.2.1.) puis en présentant les différents effets de sens (4.2.2.) et valeurs modales (4.2.3.) qui lui sont attribuées et la manière dont ceux-ci dépendent du cotexte (4.2.4.). Nous nous concentrerons également sur son emploi qualifié de post-modal (4.2.5.).

4.2.1. Le noyau sémantique de *pouvoir*

Comme susmentionné, la question de la polysémie des verbes modaux est un fait attesté depuis longtemps, en particulier pour *pouvoir* et *devoir* (Vetters & Barbet 2006 : 1). Les effets de sens et les valeurs modales de *pouvoir* sont ainsi nombreuses et découlent *a priori* d'un NS de *possibilité*, précisée *abstraite* pour De Saussure (2014b : 3-4). De Saussure (2014b) explique ainsi que tous les effets de sens et valeurs modales qu'il attribue à *pouvoir* peuvent découler de ce NS :

En ce qui concerne *pouvoir*, la permission (de faire) et l'éventualité (d'être) sont bien réductibles à des possibilités : *pouvoir* déontique introduit une nouvelle possibilité dans le monde (ou annule une impossibilité) en vertu d'un fait social qui offre une condition nécessaire. En cela, *pouvoir* déontique n'est différent de *pouvoir* radical que par l'origine de la possibilité. *Pouvoir* épistémique est une possibilité en tant que pensée (c'est une possibilité pensée qui se présente comme telle, donc une opinion subjective au sujet d'une possibilité). Ces effets de sens reposent donc sur la possibilité entendue comme compatibilité avec le monde, avec un paramétrage contextuel donnant lieu à ces effets types. (De Saussure 2014b : 8)

Ce NS a ainsi l'avantage de regrouper l'ensemble des effets de sens et valeurs modales attestées de *pouvoir*. Cependant, d'autres NS apparaissent dans la littérature à son sujet. Ainsi, pour Kronning (1996 : 33), tous les effets de sens de *pouvoir* se rattachent au concept de

³⁵ Il existe également toute une littérature sur *pouvoir* qui concerne les actes de langage, surtout indirects (Kronning 1996 : 86 ; Searle 1979 : 73, 96-97) et donc, plus généralement, son importance dans l'étude de la politesse (Chu 2008 : 145-146 ; Perret 1974 : 107). Nous nous concentrons dans cette thèse uniquement sur la sémantique des verbes modaux et n'étudierons pas leurs effets d'usage dans un cadre plus pragmatique, voire interactionnel.

problématicité. Il conceptualise cette théorie en s'inspirant de la notion des *mondes possibles* de Kratzer (1981). Cette dernière explique que « [d]ans la sémantique des mondes possibles, une proposition est examinée par rapport à l'ensemble des mondes possibles dans lesquels elle est vraie » (Kratzer 1981 : 42). L'autrice utilise cette théorie pour délimiter les verbes modaux de l'allemand et de l'anglais. En français, les verbes modaux peuvent aussi être analysés selon cette notion :

8. Pascal **peut** être à la maison.
9. Pascal **doit** être à la maison.
10. Il **faut** que Pascal rentre à la maison.
11. Pascal **veut** être à la maison.

Pour les exemples 8 et 9, la valeur modale épistémique de *pouvoir* et *devoir* montre que dans au moins un des mondes possibles, *Pascal est à la maison*. Pour *falloir* et *vouloir* – qui ne sont pas analysés avec un équivalent allemand dans la théorie de Kratzer (1981) – nous pouvons également comprendre que la notion de *vrai dans au moins un monde possible* comme étant *actualisé* ou non. Ainsi, dans l'exemple 10, le prédicat n'est pas encore actualisé et la locutrice émet donc un énoncé avec un monde dans lequel *Pascal sera à la maison s'il suit l'obligation*. Pour le verbe *vouloir* dans 11, le prédicat n'est pas actualisé non plus : on sait que Pascal n'est pas à la maison, car nous ne pouvons pas avoir la volonté de faire quelque chose qui est déjà réalisé³⁶. Selon la théorie des mondes possibles, cet énoncé implique ainsi que les désirs de Pascal sont réalisés.

Dans ces deux derniers énoncés qui ne relèvent pas de la valeur modale épistémique – et pour lesquels la théorie de mondes possibles est donc moins applicable – nous pouvons également convoquer la théorie des implicatures conversationnelles de Grice (1979)³⁷. En effet, par implicature, nous pouvons comprendre que *Pascal n'est pas (encore) à la maison* lorsque le verbe *falloir* est utilisé avec une valeur modale déontique et *vouloir* avec une valeur modale boulique.

Si nous revenons à la proposition de Kronning (1996) qui suggère que le NS de *pouvoir* est celui de la *problématicité*, cela ne semble ainsi pas justifiable. Comme nous venons de le

³⁶ Heim (in Harner 2016 : 77-79) qualifie l'équivalent anglais de *vouloir*, soit *want*, comme étant un opérateur des *mondes possibles*. Cette interprétation semble également valable pour *vouloir*.

³⁷ Nous remercions Anouch Bourmaysan pour cette suggestion qui nous a été faite lors d'une première version de cette thèse.

démontrer, cette notion de *problématicité* est également applicable à d'autres verbes modaux et ne différencie ainsi pas *pouvoir* de *devoir*, *falloir*, ou encore *vouloir*.

Gosselin (2010) propose également un autre NS que celui de *possibilité* pour *pouvoir* associant à ce verbe le NS de *compatibilité*. Selon lui, l'utilisation de ce verbe « indique [que la valeur modale exprimée par celui-ci] est *compatible* avec un ensemble de prémisses » (Gosselin 2010 : 449). Cette vision est finalement proche de celle de Kronning (1996) mentionnée à l'instant, car cela signifie que l'énoncé qui comporte *pouvoir* est compatible avec le monde de l'énoncé en question.

Ainsi, contrairement à ce que nous pourrions croire, il n'existe pas de consensus quant au NS du verbe *pouvoir*. Cela peut être dû aux nombreux effets de sens et valeurs modales listées pour ce verbe dans les analyses à son propos. Nous allons nous pencher sur celles-ci dans les deux sections suivantes.

4.2.2. Les effets de sens de *pouvoir*

Les effets de sens de *pouvoir* sont nombreux mais pas nécessairement détaillés chez les autrices qui les mentionnent. Barbet (2012), qui s'inspire des écrits de Le Querler (1996, 2001) pour lister les divers effets de sens de *pouvoir*, rappelle ainsi que ce modal peut traduire la *capacité*, la *permission* ou encore la *possibilité (matérielle)*³⁸. Pour différencier ces trois effets de sens, elle explique que « [pour la *capacité*,] l'origine de la possibilité réside dans les qualités inhérentes au sujet ; [pour la *permission*,] elle correspond à un être animé humain ou à des lois sociales ou éthiques ; [pour la *possibilité*,] elle demeure dans les conditions matérielles extérieures au sujet » (Barbet 2012 : 50)³⁹.

Desclés (2003 : 57-58) détermine les mêmes effets de sens pour *pouvoir*, mais prend comme point de départ la notion d'*instant* que l'on peut rapprocher de la théorie des *mondes possibles* de Kratzer (1981), susmentionnée dans la section précédente. Ainsi, une proposition modalisée par *pouvoir* (ou par *devoir* ou *vouloir*, les deux autres verbes modaux étudiés par Desclés (2003)), peut être *réalisée*, *non-réalisée*, *quasi-nécessaire*, *impossible*, *possible*, *probable*, ou *actualisée à l'instant t*.

³⁸ Notons ainsi que la *possibilité* semble être tant un NS qu'un effet de sens pour *pouvoir* d'après certaines autrices.

³⁹ Nous détaillerons l'influence du cotexte (notamment d'après le sujet du verbe) sur les effets de sens et valeurs modales de *pouvoir* en 4.2.4.

Sueur (1979 : 97)⁴⁰ ajoute à cette liste l'effet de sens de *non exclusion* – aussi appelée *éventualité* chez Hütsch (2020 : 22) ou encore *probabilité* chez Leeman-Bouix (2002 : 122). Les autrices ne développent pas les conditions d'emploi (ou de cotexte) nécessaires pour que *pouvoir* puisse revêtir cette interprétation, elle est simplement listée parmi d'autres.

Il existe des effets de sens plus vastes encore comme les emplois *concessifs* (12), de *minimum obligé* (13), *scalaire* (14), ou encore *sporadiques* (15 a et b). C'est De Saussure (2014b : 3-4) qui illustre ces divers effets de sens à l'aide d'exemples empruntés à d'autres autrices⁴¹ :

12. Elle **peut** pleurer, j'irai pas la voir. (Le Querler 2001 : 30)⁴²
13. Ça [le vin] **peut** être bon, à ce prix-là. (Barbet 2012b)
14. Ce que ça **peut** être ennuyeux de dîner en ville ! (Damourette & Pichon 1911-1936)
15. a) Les lions **peuvent** être dangereux. (trad. d'après Leech 1969)
b) Les Alsaciens **peuvent** être obèses. (Kleiber 1983)

Quant à Chu (2008 : 89), il distingue simplement « un *pouvoir* épistémique et un *pouvoir* de capacité » et réunit ainsi les effets de sens et les valeurs modales de ce verbe au même niveau⁴³. C'est également le cas de Rossari (2020) qui propose dans une même analyse des lectures de *permission*, *capacité* et de valeur modale épistémique. Nous allons d'ailleurs directement nous concentrer sur les valeurs modales de ce verbe qui ont été listées dans les études sur *pouvoir*.

4.2.3. Les valeurs modales de *pouvoir*

Les valeurs modales généralement attribuées à *pouvoir* sont les modalités *épistémiques*, *descriptives*, ou encore *déontiques*. La plus grande différenciation est habituellement opérée entre les valeurs modales épistémiques, et les deux autres. Ces dernières sont parfois regroupées dans une catégorie plus générale appelée *modalités radicales* (Barbet 2013). Ce sont les interprétations radicales qui comportent les effets de sens de *capacité*, de *permission* ou encore de *possibilité matérielle* (Barbet 2012 : 50). Les effets de sens les plus communément acceptés

⁴⁰ Sueur (1979 : 97) précise que dans les dictionnaires et les grammaires encore davantage d'effets de sens sont associés tant à *pouvoir* qu'à *devoir*. Il a dénombré les effets de sens suivants : « intention, faculté, capacité, possibilité, probabilité, éventualité, approximation, vraisemblance, supposition, doute, hypothèse, incertitude, etc. » (Sueur 1979 : 97). Nous reviendrons sur les points communs entre *pouvoir* et *devoir* en 4.6.1.

⁴¹ Tous les exemples qui suivent sont directement tirés de l'article de De Saussure (2014b) et les références ne figurent donc pas dans la bibliographie de ce travail mais sont à retrouver dans la publication de l'auteur.

⁴² Nous reviendrons plus en détail sur cet emploi en 4.2.5 car il peut être considéré comme post-modal.

⁴³ Ceci nous ramène à une question plus générale : est-ce vraiment nécessaire de différencier ces deux niveaux sémantiques ? Nous préférons ici distinguer les deux notions afin d'avoir une analyse plus fine encore de la sémantique des verbes modaux.

pour *pouvoir* se regroupent dans les valeurs modales *épistémiques*, considérées comme plus subjectives (Gosselin 2010 : 445) que les valeurs modales déontiques. Ces dernières, plus rarement attribuées à *pouvoir*, peuvent être relevées dans des énoncés comme :

16. **Puis**-je partir plus tôt du cours ?

L'exemple 16 dénote une demande de *permission* (Kratzer 1981 : 60), effet de sens qui est ainsi classé comme dénotant une valeur modale déontique pour *pouvoir*. Cette interprétation est inférée, car la personne à qui s'adresse la demande comprend que ce n'est pas une question de *capacité* ou de *possibilité* qui est mise en évidence ici mais bien une *permission*⁴⁴.

L'attribution d'une valeur modale ou d'un effet de sens à *pouvoir* fait toujours l'objet de beaucoup d'incertitudes. En effet, ces valeurs modales et ces effets de sens se superposent dans certains co(n)textes⁴⁵, comme le démontre notamment Le Querler (2004 : 653) avec l'énoncé suivant :

17. On **peut** partir demain.

Dans l'exemple 17, le verbe *pouvoir* peut être interprété comme une *permission* – la locutrice annonce le fait que le droit de partir lui a été accordé, ceci représente ainsi une valeur modale *déontique* – mais aussi une *possibilité matérielle* ou encore une *capacité* – les circonstances de la route, de l'emploi du temps ou d'un autre facteur font que la locutrice a l'habileté de partir le lendemain. Cette lecture s'apparente donc davantage à une modalité *épistémique* voire *descriptive*⁴⁶.

À ce jour, les diverses valeurs modales du verbe *pouvoir* n'ont pas fait l'objet de recherches approfondies, les chercheuses mettent uniquement en évidence leur existence et la difficulté de

⁴⁴ Cela rejoint également les interprétations de Searle (1972, 1979) concernant les actes indirects dénotant des demandes, ce qui démontre, une fois de plus, le lien étroit entre sémantique et pragmatique.

⁴⁵ Comme le souligne Desclés (2003), *devoir* et *vouloir* peuvent également avoir plusieurs interprétations selon le cotexte mais cette difficulté semble plus souvent soulevée pour *pouvoir*. D'ailleurs, Sitri (2013) souligne que « la quasi-totalité des études consacrées à *pouvoir* insistent sur les cas fréquents de glissement, superposition ou "indécidabilité" entre les différentes valeurs » (Sitri 2013 : 76).

⁴⁶ Notons que cet intérêt de distinguer la valeur modale d'un verbe, voire d'un énoncé, est propre aux recherches en linguistique sur la modalité. Desclés (2003 : 49) précise que les logiciennes se contentent souvent d'attribuer un seul effet de sens aux divers marqueurs modaux (qui correspondrait donc dans notre recherche au NS du verbe modal). Elles ne se focalisent ainsi pas sur toutes les interprétations possibles d'un marqueur modal. Caron et Caron-Pargue (2003 : 167) adhèrent à cette distinction, en mentionnant que le fonctionnement *linguistique* des verbes modaux n'a rien à voir avec leur fonctionnement décrit dans les travaux de *logique*. Dans ces derniers, un marqueur modal correspond généralement à une seule opération modale.

les distinguer. Toutefois, certaines autrices font mention d'éléments du cotexte pour attribuer une valeur modale ou un effet de sens plutôt qu'un autre à *pouvoir*.

4.2.4. Attribution de l'effet de sens ou de la valeur modale de *pouvoir* à l'aide du co(n)texte

Les éléments cotextuels peuvent donner des indications quant à l'interprétation sémantique de *pouvoir*. Ainsi, Sueur (1979) précise que l'effet de sens de *permission* est admis lorsque le sujet est « un [syntagme nominal] marqué [+ *humain*] » (Sueur 1979 : 99). Pour cet effet de sens, Gosselin (2010) indique que le sujet peut également représenter une institution de type « justice, morale, religion, famille, école... » (Gosselin 2010 : 445). Quant à l'effet de sens de *possibilité*, le syntagme nominal en position sujet désigne habituellement un inanimé pour Sueur (1979 : 99). Pour l'expression d'une *capacité*, celle-ci est possible lorsque la complémentation verbale désigne « une *qualité inhérente* du sujet de *pouvoir* » (Sueur 1979 : 99). L'auteur ne précise toutefois pas quel type de verbe peut désigner ce genre de caractéristique.

Notons que cette *qualité inhérente* n'est pas forcément attribuée à un sujet *humain*. Kratzer (1981 : 55-56), qui analyse l'équivalent allemand de *pouvoir* – soit *können* – montre que les sujets dénotant des objets peuvent également être définis avec des *capacités exceptionnelles*⁴⁷ comme c'est le cas du couteau dans l'exemple suivant :

18. Ce couteau **peut** découper une pierre en morceaux.

C'est lorsque les objets ont des propriétés inattendues, qu'on peut également retrouver un effet de sens de *capacité* selon l'autrice. Cela rejoint les réflexions de Van de Velde (2010) qui précise que les noms d'instruments peuvent, en position de sujet, adopter des traits d'un agent animé et auxquels on peut donc attribuer des propriétés normalement définies pour les humains. Dans leur ouvrage sur le pronom *on*, Fløttum et al. (2007) citent la recherche de Larsen (1984), qui avait également noté que « [p]our *pouvoir*, notamment la combinaison *on peut* + verbe à l'infinitif, [définit] "l'ensemble sémantique constitué par l'auxiliaire *pouvoir* et la catégorie du présent comme un *micro-contexte d'effet abstraitisant*" (Larsen 1984 : 64) » (Fløttum et al. 2007 : 22-23). Ainsi, ce ne sont pas seulement les sujets sous forme de nom qui peuvent influencer sur l'effet de sens ou la valeur modale du verbe, mais également les pronoms personnels.

⁴⁷ Ces capacités « inattendues » peuvent également s'appliquer à des sujets animés, comme le rappelle Carlier (in Sitri 2013 : 81) avec l'exemple « "Les colibris *peuvent* voler à reculons" [...] où le prédicat marque une disposition quelque peu exceptionnelle, qui est évaluée comme une performance » (Carlier in Sitri 2013 : 81).

D'autres éléments peuvent également influencer sur l'interprétation de *pouvoir*. Ainsi, Sitri (2013), dans une étude prenant en compte *pouvoir* dans un corpus de rapports éducatifs, explique que

l'interprétation sporadique s'imposera-t-elle d'autant plus que l'énoncé comporte des circonstants indiquant une itération (*dans ces moments, dans les périodes où, lorsque, dès que, régulièrement, parfois, par moments*) ou mieux encore que le prédicat avec *peut* est pris dans une alternance entre deux prédicats représentant des comportements le plus souvent antithétiques, alternance qui matérialise le caractère « sporadique » du prédicat modalisé : le balancement peut être marqué par des structures doubles comparatives (*il peut ... comme*) ou oppositives (*elle peut ... alors qu'elle peut, peut à la fois ... mais*), ou concessives (*il peut ... sans pour autant, il peut ... toutefois, il peut ... mais*), par des circonstants (*en fonction des moments, à d'autres moments*) ou des marqueurs d'addition ou d'opposition (*a contrario, aussi*). Enfin le changement, la non-stabilité du prédicat peut donner lieu à commentaire explicite (*le discours manque de cohérences, ses positions sont très fluctuantes, tient des discours changeants*). (Sitri 2013 : 80)

Dans ce même corpus, la capacité, elle, transparait grâce à

des marques indiquant une modification, un changement et avec des prédicats valorisés dans le champ social, conférant là encore à la capacité un caractère exceptionnel de réussite. Les marqueurs de changement seront des circonstanciels (du type *désormais, aujourd'hui, maintenant, au moment de l'enquête sociale, à partir de cet instant, après un échange*), des comparatifs (*plus sereinement*) ou une comparaison explicite avec un état antérieur (*ce qu'il ne parvenait pas à faire*). Quant aux prédicats, on retrouve entre autres des items analogues à ceux repérés dans les énoncés avec *pouvoir* au passé composé, en particulier pour les verbes de parole (*exprimer/évoquer ses difficultés, formuler le fait que cet hébergement lui est nécessaire / des excuses*) mais aussi des prédicats « psychologiques » (*être dans la réparation, prendre du recul face aux demandes*) ou correspondant à des actions présentées comme « difficiles à faire » pour l'enfant (*regarder un film de part en part, partager certains moments agréables avec les autres, aller chercher une éponge pour nettoyer*). (Sitri 2013 : 81)

Ces commentaires ont été approfondis par Née et al. (2014) et Née et al. (2016). Dans ces recherches, les autrices s'intéressent aux routines discursives⁴⁸. Elles relèvent que dans le genre des rapports éducatifs, *pouvoir* figure dans des patrons de type *[et] + [circonstant] + X peut/a pu + [adverbe/circonstant] + [nous] + Prédicat « difficile à dire ou à faire »* ou encore *[et] + [peut]* dans lesquels *pouvoir* exprime une *capacité* liée à un *progrès*.

⁴⁸ Selon les autrices, ces routines sont liées à des phénomènes de figement, de répétition, et s'apparentent donc aux notions de cooccurrences et de segments répétés (Née et al. 2016 : 74) Au vu des éléments qui composent la notion de routine chez les autrices, nous ne pouvons faire abstraction du fait qu'elles peuvent être rapprochées d'un stade précoce de grammaticalisation comme nous la définissons dans la section 4.1.1.

Il arrive aussi que les autrices prennent en compte une interprétation globale de l'énoncé pour accorder un effet de sens ou la valeur modale de *pouvoir*. Ainsi, Chu (2008 : 89) ajoute que

[q]uand la possibilité d'un événement est attribuée à la compétence du sujet du verbe, nous avons un *pouvoir* de capacité : quand c'est l'interlocuteur qui donne la possibilité au sujet, nous avons un *pouvoir* de permission : quand, par contre, on déduit la possibilité des circonstances, c'est un *pouvoir* de vraisemblance. (Chu 2008 : 89)

Ce linguiste ne donne toutefois pas d'exemples concrets pour chacune de ces interprétations. Sans le mentionner tel quel, Chu (2008) semble faire référence ici à un enrichissement pragmatique de l'énoncé pour l'interprétation sémantique du verbe, concept que nous avons présenté dans la section 4.1. de ce travail.

Pour résumer la potentielle influence du contexte sur l'interprétation de *pouvoir*, c'est en premier lieu le type de sujet régissant ce modal qui aide à déterminer la valeur modale ou l'effet de sens de celui-ci selon les diverses recherches que nous avons citées ici. Ces recherches contredisent les affirmations d'Abeillé et Koenig (2021a) à propos des verbes à *héritage* (chapitre 4). En effet, si ce n'est pas *pouvoir* qui sélectionne son sujet, comment ce dernier peut-il influencer l'interprétation sémantique du modal ? De plus, nous avons constaté que d'autres éléments du contexte plus large peuvent attribuer au verbe *pouvoir* un certain effet de sens ou une valeur modale. Ainsi, l'inférence obtenue à partir de l'énoncé global joue également un rôle important dans l'interprétation sémantique de *pouvoir*.

4.2.5. L'emploi post-modal de *pouvoir*

Certains emplois de *pouvoir* ne semblent pas correspondre à des catégorisations traditionnelles de la modalité, et sont ainsi qualifiés de *post-modaux* par certaines autrices. Selon Van der Auwera et Plungian (1998), la post-modalité se définit comme une désémantisation sur le plan diachronique des marqueurs de modalité qui n'exprimeraient, *a priori*, plus une valeur modale ou un effet de sens lié à leur NS, ce que nous pourrions ainsi rattacher aux premières étapes de grammaticalisation de Diewald (2002) ou Heine (2002) présentées dans la section 4.1.1. Toutefois, Van der Auwera et Plungian (1998) ne mentionnent pas la notion de grammaticalisation lorsqu'ils traitent la post-modalité dans leurs recherches.

La post-modalité concerne ainsi les occurrences des marqueurs modaux lorsque leurs NS ne sont plus perceptibles. Pour les chercheuses appliquant ce procédé à la lettre, tout écart du NS peut être considéré comme post-modal. Ainsi, Julian et Thorne (1969) expliquent que l'équivalent anglais *can* serait autre chose que modal dès qu'il ne dénote pas une *possibilité*

dans son sens le plus strict du terme. Ainsi, même une *capacité* – et donc une modalité descriptive – serait considérée comme post-modal pour ces auteurs.

Pour les chercheuses plus contemporaines et issues du monde francophone, l'emploi post-modal de *pouvoir* correspond à des usages bien plus éloignés du NS, comme c'est le cas dans l'exemple suivant⁴⁹ :

19. Elle **peut** toujours pleurer, je n'irai pas la voir.

Cette notion de post-modalité ne fait pas l'unanimité entre les chercheuses. Cet exemple 19 en particulier est analysé par De Saussure (2014a) comme étant simplement un monde possible dans lequel *elle* pleure et que ces pleurs n'ont pas la *capacité* de changer le monde :

Dans ce contexte, ouvrir la compatibilité de *pleurer* avec *ne pas aller voir*, en manifestant la conjonction de deux éventualités normalement incompatibles, implique que la relation présupposée est fausse ou ne s'applique pas dans le cas particulier : ces pleurs n'ont pas la capacité de changer le monde. Mais l'effet pragmatique ne tient qu'au fait que, dans le contexte, une hypothèse opposée existe : le locuteur est indifférent aux pleurs. Ce n'est qu'en second lieu, suggérons-nous, que l'effet d'autorisation, symbolique, peut surgir, comme dans *Ils peuvent bien écrire ce qu'ils veulent, je m'en fiche / je ne protesterai pas*. Par comparaison, un énoncé de possibilité épistémique manifeste généralement plutôt la pertinence de la possibilité, c'est pourquoi exprimer la possibilité que P communique généralement une invitation à agir en fonction de la réalisation de P : *il peut pleuvoir* incite à agir en fonction de la pluie et non en fonction du beau temps, même si la possibilité logique ne permet pas cette conclusion. Pragmatiquement, évoquer la possibilité de P implique qu'il faut se conformer à un monde où P est le cas et non à un monde où P n'est pas le cas : or, cette conclusion est évidemment inférée pragmatiquement pour assurer la pertinence de l'énoncé. (De Saussure 2014a : 121)

Le NS de la *possibilité* serait ainsi encore déductible dans un énoncé tel que celui de l'exemple 19 et il n'y aurait ainsi pas lieu de qualifier cette utilisation comme s'éloignant des catégories traditionnelles de la modalité.

Cette analyse vaut également pour les emplois rhétoriques chez Rossari (2020) qu'elle propose pour un énoncé comme celui de l'exemple suivant :

20. Je **peux** être suisse, mais je soutiens la France.

⁴⁹ Cité déjà dans l'exemple 12 dans la partie 4.2.2.

Cet emploi de *pouvoir* peut également être relié au NS de la *possibilité* mais porte sur le *dire* et non sur le *dit*. Il y a ici un effet de sens de *concession* car la locutrice peut *bien* admettre qu'elle est suisse *mais* qu'elle soutient la France.

Ainsi, les usages post-modaux du verbe *pouvoir* ne sont pas perçus comme transcendant les catégories traditionnelles de la modalité par les autrices contemporaines et francophones. Cependant, il est manifeste que certains de ces usages sont moins fréquents que d'autres, sans pour autant être complètement dissociés du NS conventionnellement attribué à *pouvoir*. Il apparaît donc difficile de parler de grammaticalisation pour ce modal, même si certaines linguistes observent une désémantisation dans certains contextes d'emploi.

4.2.6. Conclusions sur *pouvoir*

Dans cette partie, nous avons dénombré les nombreux effets de sens et valeurs modales listées dans les recherches sur *pouvoir*. Celles-ci peuvent généralement être rattachées à un NS de *possibilité*. Même si ce NS ne fait pas l'unanimité, il semble toutefois être celui qui peut regrouper le plus d'effets de sens et de valeurs modales admises pour ce verbe, même pour les emplois considérés comme post-modaux, mentionnés dans la sous-section précédente⁵⁰.

Le verbe *pouvoir* est aussi le plus fréquent des verbes que nous étudions ici⁵¹. Cette haute fréquence peut expliquer ses nombreux effets de sens et valeurs modales⁵². Généralement, on lui associe la *possibilité (matérielle)*, la *capacité*, la *permission* ou encore l'*éventualité*. Les valeurs modales que les autrices lui attribuent sont elles aussi nombreuses, allant des modalités épistémiques à déontiques, en passant par les modalités descriptives.

Même si les listes des effets de sens et des valeurs modales qui sont attribuées à *pouvoir* semblent faire consensus, il n'est pas toujours aisé de savoir quel effet ou valeur ressort dans les énoncés. Le cotexte peut aider pour attribuer tel ou tel effet de sens ou valeur modale à *pouvoir* mais les réflexions à ce sujet semblent encore trop peu nombreuses ou, du moins, peu

⁵⁰ Ce NS de la *possibilité* peut également se retrouver dans la forme impersonnelle et pronominal de *pouvoir* peu soulevée dans la littérature comme dans : « Il se *peut* qu'il soit perdu, vu qu'il est en retard. ». Ce type d'usage est davantage mentionné lors des discussions sur les formes impersonnelles, comme c'est le cas, par exemple, dans la Grande Grammaire du Français (désormais GGF), dans le chapitre des verbes impersonnels (Abeillé et al. 2021a). Cet emploi est ainsi plutôt étudié dans le cadre d'étude des valences et de la syntaxe des verbes de ce type (Abeillé & Koenig 2021a : 155) que de la sémantique des verbes modaux.

⁵¹ En effet, il apparaît toujours devant *devoir*, *falloir* et *vouloir* en termes de fréquence et ce, quel que soit le corpus consulté parmi ceux que nous utilisons pour cette étude. Nous présenterons en détail ces différents corpus dans la partie 6.5.1. de ce travail.

⁵² Pour une explication détaillée du lien entre fréquence et polysémie, voir Kuiper et al. (2018) qui, grâce à un calcul reposant sur l'indice de Zipf, prouvent que plus un mot est fréquent, plus il est polysémique.

étayées à l'aide des outils de la linguistique de corpus, permettant une objectivisation des cotextes importants dans lesquels une unité lexicale apparaît.

4.3. *Devoir*

A l'instar de *pouvoir*, *devoir* a suscité l'intérêt de beaucoup des chercheuses, notamment pour sa nature polysémique⁵³. Mais, contrairement à *pouvoir*, *devoir* possède également « un emploi de verbe lexical plein (renvoyant à la notion de dette) » (Rossari et al. 2007 : 1). Dans cette thèse, nous nous concentrerons ici uniquement sur ses emplois modaux⁵⁴.

Sur le modèle de la partie consacrée à *pouvoir*, nous allons, dans un premier temps, revenir sur le NS (4.3.1.), les valeurs modales (4.3.2.) ainsi que les effets de sens de ce verbe (4.3.3). Cependant, contrairement à *pouvoir*, nous consacrerons ici une partie de cet état de l'art aux usages grammaticalisés de *devoir* (4.3.5). Nous reviendrons aussi sur l'influence du cotexte pour l'interprétation sémantique de ce modal (4.3.6 à 4.3.8.2). Nous verrons que cette influence a été bien plus documentée pour *devoir* qu'elle ne l'a été pour *pouvoir*. Enfin, nous nous attarderons également sur les méthodes et outils employés pour étudier les caractéristiques sémantiques de ce verbe (4.3.9).

4.3.1. Le noyau sémantique de *devoir*

Le NS habituellement attribué à ce verbe est celui de la *nécessité* (Chu 2008 ; De Saussure 2014a, 2017 ; Kronning 1996 ; Rossari 2020 ; Vetters 2012), précisée parfois *abstraite*, notamment chez Vetters et Barbet (2006) afin de justifier le fait que cette *nécessité* ne s'applique pas seulement à des objets concrets.

Gosselin (2010) attribue à *devoir* un NS « moins fort » sur l'échelle de la nécessité, puisqu'il estime que *devoir* matérialise davantage la *quasi-implication*. Il « indique que la modalité sur laquelle il porte est *quasi-impliquée* par un ensemble de prémisses » (Barbet 2013 : 101), ce qui argumente, à nouveau, le lien étroit qu'unit la sémantique à la pragmatique. Avec un NS tel que celui de la *quasi-implication*, *devoir* exprimerait plus fréquemment une valeur modale

⁵³ Comme pour *pouvoir*, il existe également pour *devoir* des travaux concernant son influence dans les actes de langage directifs, résumés notamment par Kronning (1996). Nous ne reviendrons pas en détail sur ce type d'étude ici puisque nous proposons ici un panorama concernant davantage les caractéristiques sémantiques de ce verbe modal même si nous avons vu que les théories pragmatiques sont parfois nécessaires pour une étude du sens plus exhaustive.

⁵⁴ Ces deux emplois sont d'ailleurs facilement repérables car le *devoir* modal est directement suivi d'un infinitif alors que le *devoir* de dette est « suivi d'un substantif » (Huot 1974 : 13).

épistémique que déontique. Ceci démontre que, selon le NS que l'on attribue à l'entité étudiée, l'approche ne sera pas la même. La plupart des chercheuses vont ainsi chercher à voir comment la *nécessité* est exprimée à travers les différents effets de sens et valeurs modales de ce verbe selon le cotexte alors que Gosselin (2010) partira du présupposé que *devoir* est davantage un marqueur épistémique et cela influencera le reste de son analyse. Cette double perspective est intéressante car elle permet en partie de comprendre pourquoi la valeur modale épistémique est souvent attribuée à *devoir*.

4.3.2. Les valeurs modales de *devoir*

Étant donné que le NS le plus attesté pour *devoir* est celui de la *nécessité*, *devoir* sert souvent d'exemple lorsqu'il s'agit de décrire la modalité déontique. Néanmoins, il possède également des usages qui l'éloignent de celle-ci et qui peuvent ainsi être classés parmi les modalités épistémiques ou descriptives.

Comme nous venons de le constater avec Gosselin (2010) ci-dessus, il existe des autrices qui travaillent sur *devoir* en tant qu'expression de la modalité épistémique. Dans son modèle triadique, Rossari (2020) explique que la lecture épistémique de *devoir* est associée au *dire* avec « une prise en charge de l'énoncé revendiquée, [une] interprétation épistémique avec forte présomption » (Rossari 2020 : 24) contrairement à *pouvoir* qui véhicule une moins grande certitude lorsqu'il est utilisé pour exprimer cette même valeur modale. Dans les cas où *devoir* porte sur le *dit*, il aura ainsi une lecture non-épistémique, qu'elle soit *déontique* ou *descriptive*. Dans ce type d'usage, certaines chercheuses attribuent à *devoir* un effet de sens de « forte probabilité [fondée sur] l'état actuel des croyances du locuteur, qui procèdent généralement d'inférence (d'où la valeur "évidentielle" reconnue à *devoir* épistémique [...]) » (Gosselin 2010 : 440). Cette valeur « évidentielle » est entendue comme un savoir qui découle « d'opérations de création d'information » (Dendale 1994 : 25). Il est important de noter que Dendale (1994) estime que l'*évidentialité* est une valeur modale distincte de la valeur épistémique, qu'elle sert à exposer des informations *inférées*. Mais, dans les études considérant la modalité dans son sens large, le caractère « évidentiel » de *devoir* est une expression de la valeur épistémique. C'est le cas de Gosselin (2010) qui classe les *inférences* dans les « moyens d'expressions des modalités épistémiques » (Gosselin 2010 : 331) car une certaine incertitude règne autour de cette information présupposée et cette incertitude caractérise la valeur modale épistémique.

Même si les valeurs modales les plus communément attestées pour *devoir* sont les modalités *déontiques* et *épistémiques*, certaines autrices estiment que *devoir* peut également exprimer les modalités *descriptives*. C'est le cas lorsque *devoir* est utilisé pour désigner une « nécessité absolue » (Vetters & Barbet 2006 : 5) comme dans :

21. Si tu lances une pièce en l'air, elle **doit** retomber. (Exemple inspiré de Vetters (2021))

Au vu des nombreuses valeurs modales que *devoir* peut transcrire, il revêt aussi de nombreuses possibilités d'effets de sens.

4.3.3. Les effets de sens de *devoir*

Comme nous venons de l'exposer, c'est la valeur modale déontique qui est la plus communément accordée à *devoir* et ceci se remarque dans les nombreux effets de sens qui découlent directement de cette modalité.

Ainsi, les effets de sens les plus attestés sont l'*obligation* (théorique ou matérielle pour Vetters et Barbet (2006 : 3)), l'*ordre* (Desclés 2003 : 58) ou encore l'*auto-obligation* (Desclés 2003 : 58 : Vetters & Barbet 2006 : 5). Cependant, selon De Saussure (2014b : 4) l'*auto-obligation* n'est pas forcément à considérer comme dénotant une valeur modale *déontique* de *devoir* car elle est

une forme de nécessité qui n'est pas déontique au sens classique, puisque la force de l'obligation a pour origine, avec le déontique, le monde social, donc externe à l'individu [...] Une telle interprétation de *devoir* le ferait plutôt ressembler à un équivalent de l'interprétation capacitative de *pouvoir* puisque la source de la nécessité, ici, comme avec *pouvoir* de capacité, est interne au sujet [...] (De Saussure 2014b : 4).

Concernant ses valeurs modales épistémiques, nous pourrions ainsi attribuer à *devoir* l'effet de sens de *déduction* qui s'expliquerait par une « conclusion [...] nécessaire étant donné les prémisses » (De Saussure 2014a : 120). Cet effet de sens se rapproche ainsi de la définition de l'*évidentialité* donnée plus haut pour d'autres autrices ce qui explique une fois de plus pourquoi nous pouvons classer les usages *évidentiels* de *devoir* dans les classes des valeurs modales épistémiques.

Nous avons limité les effets de sens de *devoir* aux cas les plus fréquents mais il en existe encore d'autres. Kronning (1996) dénombre 37 effets de sens possibles pour *devoir* qui sont tous des nuances de *nécessité*, d'*obligation*, de *futur* ou de *probabilité*⁵⁵.

4.3.4. Une interprétation aisée des propriétés sémantiques de *devoir*

Malgré ces effets de sens et valeurs modales multiples, la plupart des autrices s'accordent à dire que nous pouvons aisément retrouver le NS de la *nécessité* dans ces usages. *Devoir* semble ainsi plus facilement interprétable que *pouvoir*⁵⁶. Chu explique notamment que :

Nous ne pensons pas qu'il s'agisse [...] des cas d'ambiguïté et pour comprendre ces énoncés, les interlocuteurs ne sont pas tenus de deviner si le *devoir* qu'il entend indique une obligation ou une probabilité. En effet, si l'on essaie de décrire la sémantique de *devoir* [...], on est tenté de dire que le verbe *devoir* exprime essentiellement l'évaluation de la part du locuteur sur la nécessité de l'évènement. Si cette nécessité est causée par une autorité, celle de la loi, de l'éthique, du bon sens etc., nous avons l'interprétation déontique de *devoir* : si cette nécessité est due à un jugement des circonstances générales, nous avons alors l'interprétation épistémique. Mais la précision de cause n'est pas nécessairement présente dans un contexte donné et dans ces cas-là une indication de la nécessité suffira pour le message. (Chu 2008 : 86)

De Saussure (2014a, 2017) conçoit également que la *nécessité* est prédominante dans tous les effets de sens et valeurs modales qu'il liste⁵⁷. Il démontre cela à l'aide de la théorie des mondes possibles (Kratzer 1981), en mentionnant que :

au moment de l'énonciation, *l'obligation de x* est compatible avec l'existence de non-*x* dans le monde, tandis que la *nécessité que x* ne l'est pas. La même observation vaut pour la nécessité épistémique (*Pierre doit être à la piscine*) qui exprime la *probabilité que x*, laquelle est compatible avec non-*x* : ces deux modalités produisent des assertions plus "faibles" que la nécessité. (De Saussure 2014a : 119)

⁵⁵ Les effets de sens qu'il cite sont les suivants : la *nécessité*, la *nécessité inéluctable*, l'*obligation*, l'*obligation morale*, l'*obligation sociale*, l'*obligation morale atténuée*, l'*obligation matérielle*, l'*ordre*, la *suggestion*, le *reproche*, le *regret*, l'*interdiction*, l'*intention*, le *convenu*, la *convenance*, la *destination*, la *délibération*, la *destinée*, le *futur*, le *futur proche*, le *futur indéfini*, le *futur virtuel*, le *futur probable*, le *futur d'obligation*, l'*irréel du passé*, le *futur du passé subjectif*, le *futur du passé 'des journalistes'*, le *futur du passé 'des historiens'*, l'*affirmation atténuée*, la *possibilité*, l'*obligation logique*, l'*éventualité*, la *supposition*, la *probabilité*, la *vraisemblance*, l'*estimation approximative* et l'*hypothèse*.

⁵⁶ Rappelons notamment l'exemple 17 cité plus haut, dans lequel *pouvoir* pouvait être interprété comme une permission ou une capacité sans qu'un effet de sens ne prévale sur l'autre.

⁵⁷ « [...] deontic necessity; [...] alethic necessity; [...] epistemic necessity; [...] temporal necessity (what was to happen); [...] internal necessity [...] » (De Saussure 2017 : 283). Relevons que De Saussure (2017) ne semble pas différencier les effets de sens et des valeurs modales comme nous le proposons dans cette thèse.

L'assertion « forte » du NS de la nécessité de *devoir* pourrait ainsi expliquer pourquoi celui-ci s'interprète plus facilement que son pendant de la *possibilité*, soit *pouvoir*.

4.3.5. La grammaticalisation de *devoir*

Contrairement à *pouvoir*, la notion de grammaticalisation est appliquée à *devoir* pour certains de ses usages. C'est le cas notamment lorsqu'il est classé en tant qu'auxiliaire et donc considéré comme un élément plus grammatical que sémantique dans une définition stricte. Ceci est notamment le cas pour l'effet de sens de *probabilité*, qui s'assimile selon certaines autrices à une fonction d'auxiliaire de futur (De Saussure 2014b : 4), dénotant plus précisément un « futur proche » (Leeman-Bouix 2002 : 122) voire un « futur de programmation » (Larrea 2004 : 749). Dans de telles situations, *devoir* ne serait ainsi plus un verbe plein – exprimant son NS de *nécessité* ou de *quasi-implication* – mais « un auxiliaire, dépourvu de sémantisme propre, et dont le rôle serait seulement de modaliser l'énoncé dans lequel il apparaît (probabilité, incertitude liée au futur) » (Barbet 2013 : 97). Puisque le NS de *nécessité* reste récupérable dans des exemples de ce type, nous nous situons dans les premières phases de grammaticalisation, que ce soit celles mentionnées chez Heine (2002) ou chez Diewald (2002).

En revanche, le processus semble davantage entamé pour la locution *je dois dire que* + *p*. Kronning (1988) – résumé par Pusch (2007 : 36) – considère cette formule comme un « indicateur pragmatique parenthétique transparent », voire même un « connecteur concessif » qui lie « un énoncé *q* à un énoncé précédent *p* en établissant une relation de concessivité entre les deux » (Pusch 2007 : 37). En ce sens, il peut même être compris comme un MD selon les critères que nous avons soulevés de Dostie (2004). Cependant, même dans des usages tels que celui-ci, nous nous pouvons admettre que le NS de *devoir* est encore récupérable. Avec *je dois dire que*, nous pouvons comprendre qu'il est *nécessaire* de dire *p* compte tenu de divers éléments dont nous avons connaissance⁵⁸.

4.3.6. Interprétation des effets de sens et des valeurs modales de *devoir* grâce au sujet grammatical

Comme pour *pouvoir*, les effets de sens et valeurs modales de *devoir* dépendent du cotexte et des autrices ont dénombré quelques régularités à ce propos. Huot (1974) élabore tout un modèle des effets de sens de *devoir* en se basant non seulement sur les variations morphologiques –

⁵⁸ Pusch (2007) propose cette analyse pour *il faut dire que* (4.4.5.3.).

donc selon le temps et le mode verbal auquel *devoir* est conjugué – mais aussi sur les « rapports sémantiques » (Huot 1974 : 17) que le modal entretient avec les éléments qui l’entourent. Elle analyse notamment l’influence des déterminants qui précèdent les noms en position sujet. Ainsi, les articles *le* et *un*, lorsqu’ils se trouvent devant « un nom désignant une classe entière » rendent l’interprétation de *devoir* « plus spontanément⁵⁹ [...] comme un verbe exprimant l’obligation » (Huot 1974 : 32). Pour les déterminants quantifieurs comme avec *tout (tous)*, *chaque*, *chacun*, *nul*, *pas un*, *aucun*, *personne*, *rien*, *devoir* « exprime toujours l’obligation » mais varie plus souvent avec *plusieurs*, *quelques*, *beaucoup (de)*, *peu (de)* même si « en dehors d’un contexte précis, la valeur de probabilité est celle qui s’impose le plus spontanément » (Huot 1974 : 33). C’est également Huot (1974) qui propose un inventaire des effets de sens de *devoir* selon le type de pronom personnel qui le régit. Ainsi, les pronoms des premières et deuxième personnes – *je*, *tu*, *nous*, *vous* – « suffisent à faire de DEVOIR un verbe ayant une valeur d’obligation » (Huot 1974 : 34). Quant aux pronoms *il/ils/elle/elles*, ceux-ci n’exercent « aucune influence particulière sur DEVOIR » sauf quand « IL est employé comme impersonnel, c’est-à-dire où il n’est le substitut d’aucun SN, [dans ce cas] DEVOIR n’a qu’une seule valeur, celle de probabilité (cf. Dubois, 1969 a : 119) [...] » (Huot 1974 : 40).

A propos de *on*, Huot rappelle que

Dubois (1965 : 111 sq) a fait remarquer que ON, “arbitrairement rangé parmi les indéfinis devait être intégré aux pronoms personnels”. Par ailleurs, il a souligné que ce pronom “sans aucune marque spécifique de personne était susceptible de se substituer à tous les autres pronoms personnels”. Dans la mesure où certains de ces pronoms personnels exercent une influence sur la valeur de DEVOIR, ON, substitut de ces pronoms, exercera la même influence sur DEVOIR [...] (Huot 1974 : 40)

Vetters et Barbet (2006) prennent également en compte le sujet pour l’interprétation de *devoir*. Elles expliquent, par exemple, que la différence entre un effet de sens d’*obligation théorique* et d’*obligation matérielle* repose sur la source de la nécessité. Dans le premier cas, il s’agit d’ « une personne ou une loi », ou encore de la « morale » ou de la « religion » (Vetters 2012 : 39) qui émet l’obligation et dans le second, « la source réside dans des circonstances matérielles » (Vetters & Barbet 2006 : 3). Il est cependant important de distinguer la source de la nécessité et le sujet grammatical du verbe⁶⁰. Ceux-ci ne coïncident pas toujours pour *devoir* :

⁵⁹ Notons qu’il n’y a pas de généralisation absolue dans les dires de Huot (1974) seulement une haute probabilité qu’une interprétation soit plus plausible selon le type d’éléments qui précèdent ou suivent *devoir*.

⁶⁰ Nous pourrions ici faire la différence entre *sujet syntaxique* – donc le nom ou le pronom avec lequel s’accorde le verbe et qui se trouve généralement antéposé au verbe – et *sujet sémantique*, qui survient souvent après le verbe, notamment dans les tournures passives, et qui est ainsi la personne ou l’objet qui réalise l’action dénotée par le verbe sans se retrouver en position de sujet syntaxique (Girard 2004 : 43).

22. Je **dois** ranger ma chambre, ma maman me l'a demandé.

Dans cet exemple, nous pouvons bien constater que ce n'est pas *je* qui est à l'origine de l'obligation, mais bien la *maman* de *je*. Sueur (1979) semble faire une sorte d'amalgame entre ces deux notions lorsqu'il écrit que la source de l'obligation est un être animé dans l'exemple suivant⁶¹ :

23. Pierre **doit** venir.

Dans 23, *devoir* exprimerait ainsi un effet de sens d'*obligation*. Sueur (1979) distingue cet effet de sens de celui de *nécessité* par une différenciation du sujet grammatical : dans ces cas, le sujet serait doté du trait *inanimé*. En ceci, les autrices *infèrent* ici la source de l'obligation sans que celle-ci ne soit nécessairement représentée par le sujet grammatical.

Kronning (1996), quant à lui, distingue surtout les valeurs déontiques et épistémiques de *devoir*. Ainsi, lorsque le sujet n'est pas agentif, *devoir* aura ainsi une portée *de dicto*, paraphrasable par *il est obligatoire de*.

Les autrices ont donc analysé les effets de sens et valeurs modales de *devoir* à travers de nombreux paramètres du sujet grammatical, que ce soient des traits sémantiques – à travers le trait animé ou inanimé du nom qui précède le nom –, des paramètres grammaticaux – comme le type de déterminant –, ou encore les pronoms personnels.

4.3.7. Interprétation des effets de sens et des valeurs modales de *devoir* grâce à la complémentation verbale sous forme d'un verbe à l'infinitif

Il existe également des analyses qui se focalisent sur le complément verbal sous forme de verbe à l'infinitif pour déterminer l'effet de sens ou la valeur modale de *devoir*. Huot (1974), auparavant citée, s'est focalisée sur ce paramètre-là. Elle explique notamment que l'effet de sens de *futur* pour *devoir* est possible lorsqu'il est suivi de verbes de mouvement⁶² (Huot 1974 : 60-62). Lorsqu'il est suivi d'un verbe d'état ou du verbe *être*, « il exprime alors la probabilité [...] » (Huot 1974 : 63).

⁶¹ Notons toutefois qu'il ne semble pas complètement convaincu par sa terminologie étant donné qu'il met le mot « source » entre guillemets et que plus loin dans le texte il parle de « sujet », dans le sens de « sujet grammatical ».

⁶² L'autrice ne donne que trois exemples de verbes pour cette catégorie, soit *aller*, *venir* et *partir*.

Chu (2008 : 84), en reprenant les études de Huot (1974), précise que si le verbe qui suit *devoir* est impersonnel, *devoir* aura un effet de sens de *probabilité* (24) mais exprimera une *obligation* (25) lorsqu'il est suivi d'un autre verbe modal⁶³ :

24. Il **doit** pleuvoir.⁶⁴

25. Jean va **devoir** faire ce travail.

Il est assez étonnant de constater que l'exemple 25 montre une utilisation de *devoir* à l'infinitif, faisant ainsi partie d'une chaîne verbale plus grande⁶⁵. Afin de corroborer ses réflexions, Chu précise que pour reconnaître un effet de sens d'*obligation*, il est possible de passer par un test de pronominalisation : « Le *devoir* d'obligation peut avoir le pronom clitique *le* comme complément, tandis que le *devoir* de probabilité ne le peut pas. Ex. *Elle le doit.* (nécessairement *devoir* = "obligation") [...] » (Chu 2008 : 84)⁶⁶.

Les réflexions concernant l'interprétation de *devoir* selon le verbe qui suit sont relativement plus importantes que pour *pouvoir* mais restent moins nombreuses que les observations concernant l'influence du sujet.

4.3.8. Interprétation des effets de sens et valeurs modales de *devoir* selon d'autres paramètres

En plus du cotexte proche et obligatoire de *devoir* – le sujet grammatical et la complémentation verbale –, d'autres facteurs sont pris en compte dans certaines analyses des effets de sens et valeurs modales de ce verbe. Cela est le cas notamment de la *négation*, mais également du temps et mode verbal auquel *devoir* est conjugué.

⁶³ Comme nous le mentionnions dans le chapitre 2. de ce travail, Chu (2008) a une vision très large de la modalité c'est pourquoi *faire* est considéré ici comme un modal.

⁶⁴ Cependant, comme nous l'a suggéré Marion Bendinelli, cette interprétation n'est pas automatique. En effet, nous pouvons imaginer qu'il n'a pas plu depuis des semaines, et que les agricultrices risquent de perdre leurs cultures. L'une d'entre elles pourrait dire « Il **doit** pleuvoir », car il n'y a pas d'autre alternative, il est nécessaire qu'il pleuve. La valeur modale déontique serait donc ici également possible selon le co(n)texte d'énonciation.

⁶⁵ Cela peut rappeler l'étude de Sueur (1977) qui lui aussi avait analysé *pouvoir* et *devoir* par rapport aux restrictions des verbes qui pouvaient se placer avant ceux-ci lorsqu'ils étaient employés à l'infinitif et donc eux-mêmes modalisés.

⁶⁶ Ce test a été initialement pensé par Huot (1974) et a été reformulé par Chu (2008) dans son ouvrage.

4.3.8.1. La négation

Pour *devoir*, la négation peut exercer une influence sur l'effet de sens de ce modal⁶⁷. Ainsi, *ne pas devoir* traduirait un effet de sens de *possibilité* (Sueur 1977 : 382) que Gosselin qualifie de *facultatif*, donc « le fait de ne pas être obligé de faire quelque chose » (Gosselin 2010 : 84-85). Il remet toutefois vite cette terminologie en cause lorsqu'il précise que « cette expression marque ordinairement – en dépit de ce qui était logiquement prévisible – l'interdit : “*Vous ne devez pas fumer*” n'équivaut nullement à “*Vous n'êtes pas obligé de fumer*” » (Gosselin 2010 : 84-85). Il est aussi important de noter que cet effet de sens d'*interdiction* chez Gosselin (2010 : 84) peut être traduit via une négation tant avant ou après le modal comme dans *il doit ne pas fumer*.

4.3.8.2. Le temps verbal

En plus du cotexte, d'autres paramètres peuvent donner des indices sur l'effet de sens de *devoir*. A cet égard, nous pouvons évoquer le temps verbal auquel *devoir* est conjugué. A titre d'exemples, il y a notamment le temps du futur de l'indicatif, qui « donne à DEVOIR une valeur d'obligation, quelle que soit la forme de l'infinitif qui suit. » (Huot 1974 : 34 et 39). Dendale (1994), lui, mentionne le fait que l'effet de sens de *probabilité* peut être véhiculé par le modal, qu'il soit conjugué avec un temps du passé, du présent ou du futur (Dendale 1994 : 24). Le temps verbal n'aurait ainsi selon lui pas d'importance pour l'expression de cet effet de sens en particulier.

Rossari et al. (2018a) compare l'emploi du conditionnel présent au présent de l'indicatif pour *devoir*. Elles expliquent que le conditionnel semble restreindre la palette des valeurs que le modal peut prendre, se limitant principalement à un effet de sens de prévision. En revanche, au présent, *devoir* peut revêtir des effets de sens et valeur modales plus nombreuses, notamment de *nécessité/obligation*, de futur et de valeur modale épistémique.

Ces remarques de l'influence du temps verbal sur l'effet de sens de *devoir* ne sont toutefois pas si abondantes, ni aussi déterminantes que cela. En effet, « [o]n constate que seuls quelques temps (peu nombreux) exercent une influence très nette sur la valeur de DEVOIR. Aux autres temps, la valeur de DEVOIR dépend d'autres éléments de la phrase [...] » (Huot 1974 : 48).

⁶⁷ C'est également le cas pour *pouvoir* dans une moindre mesure. Ainsi, une *permission* niée serait une *obligation de ne pas faire* (Narrog 2012b : 188) voire une *contrainte* (Sueur 1977 : 382). *Pouvoir* s'interpréterait ainsi dans ce type d'énoncé avec une valeur modale déontique ou encore descriptive.

Les éléments co(n)textuels semblent ainsi plus importants pour interpréter les caractéristiques sémantiques du modal.

4.3.9. Les méthodologies utilisées pour étudier *devoir*

Contrairement à *pouvoir*, *devoir* a été étudié avec des outils quantitatifs. Hütsch (2020) souligne cette particularité en écrivant que « [p]our les verbes modaux français, il semble que seul *devoir* ait été analysé d'un point de vue quantitatif »⁶⁸ (Hütsch 2020 : 75). Elle cite notamment l'étude de Kronning (1996) qui « a analysé un corpus composé de romans, pièces de théâtre, ouvrages historiques et journaux, qui totalise 3339 occurrences de *devoir* » (Hütsch 2020 : 75). Le but de Kronning (1996) était de démontrer quelles étaient les valeurs modales les plus souvent exprimées par ce verbe. Ainsi,

[s]elon les résultats de Kronning (1996 : 139), environ un quart des occurrences de *devoir* relèvent de l'emploi épistémique de ce verbe et les trois-quarts restants des emplois déontique et aléthique. Sans fournir de répartition quantitative exacte entre ces deux derniers, l'auteur postule que c'est l'emploi déontique le plus fréquent des trois emplois qu'il différencie pour *devoir*. (Hütsch 2020 : 75)

Pour son époque, Kronning (1996) a été donc un précurseur. En effet, « aucune étude consacrée au verbe *devoir* ne repose [...] sur un corpus systématiquement dépouillé de cette envergure » (Kronning 1996 : 138).

Il est intéressant de constater que cette étude a traité de manière quantitative des occurrences de *devoir* afin de délimiter le pourcentage de représentation de chacune de ses valeurs modales. Cependant, Kronning (1996) n'utilise ici aucun outil statistique de la linguistique de corpus en tant que tel puisqu'il classe simplement des exemples authentiques dans diverses catégories. La même remarque vaut pour les études de Chu (2008) ou de Huot (1974) qui utilisent, certes, des corpus pour relever des exemples authentiques mais qui n'effectuent aucune analyse quantitative à partir de ces données⁶⁹. Ce type d'approche – qui peut tout de même s'insérer dans la méthodologie de la linguistique de corpus – est plus qualitative que celle que nous mettons en application dans cette thèse. En effet, cette approche qualitative des corpus

⁶⁸ Nous reviendrons cependant sur des analyses quantitatives pour *falloir* et *vouloir*, deux modaux que Hütsch (2020) n'a pas pris en compte dans son analyse et qui ont donc échappé à son commentaire.

⁶⁹ Notons toutefois que Huot (1974) a le mérite de proposer des descriptions très précises sur le profil combinatoire de *devoir* comme nous l'avons démontré tout au long de ce chapitre. Même si son analyse ne repose pas sur des données statistiques en tant que telles, elle est très méthodique et minutieuse concernant les divers tableaux qu'elle élabore à l'aide d'exemples authentiques tirés d'un corpus journalistique.

est focalisée, en ce sens qu'elle implique plus ou moins le fait que le chercheur a déjà en tête une hypothèse linguistique avant d'aller regarder ce qu'il en est en corpus. Elle sous-entend également que le résultat obtenu suffira à confirmer ou à infirmer l'hypothèse, et qu'en tout cas l'exploitation et l'exploration du corpus ne vont pas plus loin : l'extraction d'exemples authentiques se suffit à elle-même. (Poudat & Landragin 2017 : 31)

Il nous semble donc que Hütsch (2020) et, en général, les travaux de la Chaire de Linguistique française de l'Université de Neuchâtel⁷⁰, soient les seuls à appliquer une analyse quantitative de la linguistique de corpus pour l'étude des verbes modaux en français et plus précisément pour *devoir*.

4.3.10. Conclusions sur *devoir*

Si l'on résume les aspects sémantiques en lien avec *devoir*, nous savons que généralement, on attribue à ce verbe un NS de *nécessité*, et que cette *nécessité* peut se traduire en différents effets de sens et valeurs modales, comme le résume De Saussure (2017) :

En bref, la nécessité épistémique avec *devoir* [...] peut se traduire comme : en considérant toutes mes connaissances sur la question, *je conclus nécessairement que P*. Mais mes connaissances peuvent être plus ou moins certaines. Si elles sont certaines, la nécessité épistémique sera semblable à une nécessité matérielle, comme dans « tout homme doit mourir » ou même une nécessité logico-aléthique, comme dans « deux plus deux doivent faire quatre » (l'idée que la nécessité aléthique est une version spécifique de la nécessité épistémique est suggérée par Lyons en 1977). Quant à la nécessité déontique, une façon de l'appréhender en termes de nécessité est d'introduire des conditions nécessaires : la nécessité déontique sera alors traitée comme l'expression de conditions nécessaires pour atteindre certains objectifs ou pour correspondre à des états du monde souhaitables. (De Saussure 2017 : 291)

Nous avons pu démontrer que les effets de sens de *devoir* étaient très nombreux, sûrement plus nombreux que pour les autres verbes que nous analysons dans cette thèse. De plus, tout comme *pouvoir*, le cotexte du verbe peut nous donner des indices sur l'effet de sens ou la valeur modale exprimée. Retenons également que *devoir* se grammaticalise, en particulier dans la locution *je dois dire que*. Ce modal a aussi le mérite d'avoir été étudié avec des outils de la linguistique de corpus, même si certaines de ces analyses se contentent d'utiliser les corpus pour relever des exemples authentiques.

⁷⁰ Pour des études quantitatives sur *devoir* nous pouvons ainsi citer Rossari (2020) ; Rossari et al. (2020) ; Rossari et al. (2007) ; Rossari et al. (2018a, 2018b) ; Rossari et al. (2016).

4.4. *Falloir*

Falloir est le troisième verbe que nous considérons dans notre thèse. Il est nettement moins étudié que *pouvoir* et *devoir* même s'il est généralement classé aux côtés de ce dernier dans les modalités déontiques.

Ce manque d'intérêt pour *falloir* dans le cadre des études sur les marqueurs modaux peut être dû au fait qu'il s'agisse d'un verbe impersonnel. En effet, *falloir* semble plus souvent mentionné pour ce trait syntaxique que pour son caractère modal. Dans la GGF, par exemple, au chapitre consacré aux valences des verbes, il est cité parmi les formes prototypiques des verbes impersonnels (Abeillé & Koenig 2021a : 155) mais également de ceux nécessitant un complément à l'infinitif, à l'instar des autres modaux étudiés ici (Abeillé & Koenig 2021a : 167). Cette caractéristique est aussi mise en avant par Vetters (2021) dans le chapitre des verbes modaux dans la GGF également. Il classe *falloir* dans les constructions impersonnelles de la nécessité comme *il est nécessaire* ou *il est obligatoire* et non pas aux côtés des verbes comme *devoir*, *avoir à* ou *être à* (Vetters 2021 : 1349).

Dans cette partie, nous allons traiter, à l'instar de Valiukienė (2022), « le fonctionnement sémantique du verbe français *falloir* et analyser les fonctions des unités comportant ce verbe en tant que composante » (Valiukienė 2022 : 2). Pour ce faire, nous passerons en revue, comme pour les deux premiers verbes de cette étude, ce qui est mentionné dans les études au sujet de son NS (4.4.2.) ainsi que de ses effets de sens (4.4.3.) et valeurs modales (4.4.4.). Afin de capturer au mieux ses propriétés sémantiques, nous examinerons également les études qui se sont consacrées à l'étymologie de ce verbe (4.4.1.), mais aussi à celles qui ont soulevé ses emplois grammaticalisés ou en voie de grammaticalisation (4.4.5.). Pour finir, nous nous pencherons sur les méthodologies que les chercheuses ont déployées jusqu'à présent pour en révéler les caractéristiques sémantiques (4.4.6.).

4.4.1. L'étymologie et l'évolution du noyau sémantique de *falloir* à travers le temps

La recherche étymologique sur le verbe *falloir* a suscité l'intérêt des linguistes, qui se sont attelées à déterminer les mécanismes par lesquels la notion de *nécessité* a été exprimée à travers un verbe impersonnel. L'origine de ce modal a été remontée jusqu'au latin. En effet, ce verbe

vient du latin classique *fallere*, “tromper ou manquer⁷¹”, devenu *fallire* en latin vulgaire. Un changement de radical est intervenu au XII^e siècle à la 1^{ère} personne qui est devenue *je fail*. Les autres formes, à savoir – *tu faus* : *il, elle faut* : *nous falons* : *vous falez* : *ils, elles falent* – étaient aussi attestées. (Valiukienė 2022 : 4)

Ce n’est qu’à partir du XIV^e siècle (Herslund 2003 : 69) que son emploi impersonnel a été attesté et c’est celui-ci qui, aujourd’hui, a une lecture de valeur modale déontique (Herslund 2003 : 71). En effet, selon Herslund (2003), nous pourrions ainsi paraphraser cette évolution de la sorte :

Topos : « Quand il y a un trou, il faut le combler » (déontique)
 « *x fait défaut à y* » → « *y doit (faire) x* »
faillir → *falloir*

Figure 2 : Explication du passage de la lecture du « manque » à la lecture de la valeur modale déontique de *falloir* dans Herslund (2003)

Le manque de quelque chose expliquerait ainsi la *nécessité* de le combler exprimée par *falloir*. L’étude de Pusch (2007) complète ces remarques étymologiques en expliquant que

[c]et usage modal impersonnel du verbe *falloir* en français moderne est le stade final d’un changement sémantique et syntaxique (actanciel) entamé en ancien français où le verbe *faillir* signifiait “manquer à, faire défaut”, mais aussi “s’arrêter” et “manquer = ne pas exister”. *Faillir* codait l’objet du manque comme sujet, tandis que le verbe *falloir* moderne code l’objet de la nécessité comme complément d’objet direct [...] (Pusch 2007 : 32)

Valiukienė (2022) rappelle aussi qu’ « [e]n français moderne le verbe a gardé son sens canonique – *faire besoin, faire défaut, manquer* – [lorsqu’il] est suivi d’un syntagme nominal [...] » (Valiukienė 2022 : 4). Les traces de ses plus anciennes significations sont donc encore visibles aujourd’hui.

Quelle que soit sa construction syntaxique (suivi d’un syntagme nominal ou verbal), nous pouvons ainsi comprendre ce verbe comme exprimant le fait que si *quelque chose fait défaut* il est alors *nécessaire de le combler*⁷². Cette *nécessité* est devenue externe à l’actant à travers le temps et ceci s’est manifesté par un emploi impersonnel attesté depuis sept siècles maintenant. Ce sème de *nécessité* – qui fait partie des notions phares de la modalité – explique ainsi

⁷¹ Dans une version antérieure de cette thèse, Francesca Dell’Oro nous a fait remarquer que le verbe *falloir* avait à l’origine davantage le sens de *manquer à sa parole* plutôt que simplement *manquer*. Ainsi, il signifiait *faillir à ce qu’on avait promis de faire*, un sens que nous ne retrouvons plus dans l’usage actuel de *falloir*, mais qui demeure encore dans *faillir*. Le verbe *faillir* représenterait ainsi un stade précoce du développement de *falloir* comme le précise le dictionnaire électronique Antidote (2018). Nous la remercions pour cette précision.

⁷² Il peut être pertinent de noter que l’équivalent italien de *falloir*, soit le verbe impersonnel *occorre* est, quant à lui, dérivé de verbes de mouvement (Dell’Oro 2022) et que l’étymologie semble ainsi propre au français malgré sa proximité avec une autre langue romane.

l'intégration de *falloir* dans le paradigme des verbes modaux malgré sa nature syntaxique particulière.

4.4.2. Le noyau sémantique de *falloir*

A priori, il n'y a pas de raison d'accorder à *falloir* un autre NS que celui de *devoir* à savoir celui de *nécessité*. C'est d'ailleurs cette lecture que propose Rossari (2020) dans son modèle triadique. Dans leur grammaire, Riegel et al. (2008) soulignent également que « *falloir* est à rapprocher, pour le sens, de *devoir* » (Riegel et al. 2008 : 285).

Or, Caron et Caron-Pargue (2003) apportent une nuance à ce propos lorsqu'elles comparent *falloir* aux côtés de *devoir* et *pouvoir*. Ainsi, elles attestent le NS de *possibilité* et de *nécessité* pour les deux verbes que nous avons déjà traités, mais elles considèrent que *falloir* exprime davantage la notion de *contrainte* : « Ce qui ressort des données, c'est que *falloir* exprime la perception d'une contrainte : un événement, une activité ou un état donné est déterminé par autre chose, qui peut être une condition explicitement formulée ou le cours général du monde » (Caron & Caron-Pargue 2003 : 167). Même si les autrices ne le mentionnent pas, nous comprenons que l'interprétation du NS de *contrainte* associé à *falloir* pourrait découler de sa nature de verbe impersonnel. Ainsi, une *contrainte* semble représenter une *obligation* provenant d'une source plus « externe » et « générale » que la *nécessité*, dont la source est plus souvent explicitement mentionnée. Contrairement à la *nécessité*, qui peut être interne à la participante, la *contrainte* semble toujours être externe à celle-ci. C'est donc en ceci que les NS de *devoir* et *falloir* se différencieraient.

Il est vrai qu'en dehors de cette étude de Caron et Caron-Pargue (2003), le NS de *falloir* n'est pas mentionné explicitement par les chercheuses. Il ne semble ainsi pas y avoir de consensus affirmé autour de cette détermination du NS de *contrainte*. Nous pourrions tout aussi bien admettre un NS de la *nécessité* pour ce verbe puisqu'il est classé, aux côtés de *devoir*, comme un marqueur de la valeur modale déontique. Dans cette thèse, nous emploierons indifféremment les deux notions pour parler du NS de *falloir* puisqu'elles sont très proches. D'ailleurs, la nuance entre ces deux termes est tellement fine qu'elle n'influe pas sur les effets de sens et les valeurs modales qui ressortent pour *falloir* et qui sont similaires à celles relevées pour *devoir*.

4.4.3. Les effets de sens de *falloir*

Que le NS de *falloir* soit interprété comme exprimant une *contrainte*, comme exposé par Caron et Caron-Pargue (2003), ou comme une *nécessité*, en raison de sa proximité sémantique avec

devoir, les effets de sens qui en découlent demeurent les mêmes. C'est pourquoi nous considérons ici que les deux NS sont ici acceptables pour *falloir*.

Caron et Caron-Pargue (2003) ont attribué à *falloir* quatre effets de sens avec des nuances à peine perceptibles. Ainsi, elles exposent comme effets de sens une *contrainte fondée sur un objectif ou un but* (26), une *contrainte fondée sur une situation* (27), une *nécessité* (28) ou encore une *contrainte liée au destin* (29) :

- 26. Pour que le vin soit bon, il ***faut*** un été sec.
- 27. Il ***faut*** s'arrêter de travailler, car il est tard.
- 28. Il ***faut*** accepter ce qu'on ne peut éviter.
- 29. Il ***faut*** qu'on me téléphone quand je suis occupée !⁷³

Selon elles, les *contraintes* peuvent donc être *spécifiques* (26 et 27), *générales* (28 et 29), *comprises sur la base de facteurs rationnels* (28) ou *incompréhensibles car irrationnelles* (29), *subjectives* (26) ou encore *objectives* (27).

Ces catégories ne nous semblent pas complètement imperméables. Les nuances entre celles-ci ne sont pas documentées ni précisées dans la recherche de Caron et Caron-Pargue (2003) et sont donc plus que contestables. Par exemple, l'énoncé suivant est classé comme étant un effet de sens de *nécessité* par les autrices, alors qu'il pourrait également être de l'ordre de l'*hypothèse* et donc véhiculer une valeur modale épistémique :

- 30. Pour dire des choses pareilles, il ***faut*** être fou.

En effet, Vetters (2021) illustre la valeur modale épistémique de *falloir* avec l'exemple suivant :

- 31. Il ***faut*** que Pierre ait perdu la tête pour faire un tel choix.

L'interprétation des effets de sens de *falloir* ne fait donc pas consensus. Valiukienė (2022 : 8) se rapproche plutôt de l'interprétation de Vetters (2021), car elle attribue à ce type d'usage l'étiquette d'*emplois subjectifs*. Dans ces cas, *falloir* exprime un effet de sens de *quasi-certitude* :

- 32. Il ***faut*** être aveugle pour ne pas voir un signe aussi manifeste ! (Valiukienė 2022 : 8)

⁷³ Tous les exemples sont de Caron et Caron-Pargue (2003).

Selon l'autrice, cette *quasi-certitude* pourrait également être de l'ordre d'une conjecture, qui « est fondée sur la probabilité ou la plausibilité qui présuppose une argumentation – [...] par induction ou par déduction, – mettant en jeu des prémisses et une conclusion » (Valiukienė 2022 : 7). Larreya (2004) propose une dénomination similaire en classant *falloir* dans ce type d'usage parmi les valeurs modales impliquées, à rapprocher donc de la notion d'évidentialité telle que définie dans la section 4.3.2. de ce travail. Dans une terminologie plus traditionnelle des classes modales, nous pouvons considérer ces utilisations de *falloir* comme relevant d'une valeur modale épistémique, puisqu'elles se rapprochent de l'usage que décrit Vetters (2021) pour l'exemple 32.

Parmi les autres effets de sens subjectifs mentionnés par Valiukienė (2022 : 7) figurent des emplois plus oralisés, avec le pronom défectif :

33. **Faut** se voir un jour ! (Valiukienė 2022 : 7)

Cet emploi peut véhiculer un effet de sens de *souhait* que Valiukienė (2022 : 7) tire des entrées du TLFi⁷⁴ pour *falloir*. Cet effet de sens est majoritairement associé à des constructions de *falloir* avec un SN :

34. C'est une maison comme ça qu'il me **faudrait** ! (Valiukienė 2022 : 7)

Dans les effets de sens plus proches du NS de *falloir*, nous retrouvons celui d'*obligation*. Comme pour *devoir*, cette *obligation* peut varier d'intensité et « aller de l'obligation rigoureuse à l'obligation faible, aux énoncés d'ordre injonctif aux conseils atténués » (Valiukienė 2022 : 7). Il existe d'ailleurs des chercheuses qui se concentrent uniquement sur ce type d'usage de *falloir* à l'instar de Patard (2015) qui considère ce verbe comme un marqueur d'*obligation morale* (Patard 2015 : 148) sans qu'aucune autre valeur modale ne soit envisagée.

Il est intéressant de relever que les remarques sur les divers effets de sens relevés pour *falloir* sont issues de recherches plutôt récentes. Longtemps comparé à *devoir*, nous pouvons ici voir que « *falloir*, qui serait logiquement équivalent à *devoir*, montre un schéma de significations complètement différent : il établit une relation de dépendance entre l'état ou l'événement décrit et les conditions dans lesquelles il se produit » (Caron & Caron-Pargue 2003 : 167). Ceci est donc surtout le cas pour les effets de sens qui s'éloignent du NS comme nous l'avons vu pour les exemples 30 à 34. Mentionnons également que ces études récentes se concentrent sur les

⁷⁴ Trésor de la Langue Française informatisé : <http://atilf.atilf.fr/> (consulté le 03.07.2024).

effets de sens et valeurs modales moins souvent attestées pour *falloir* – qui a longtemps été décrit comme uniquement un verbe de la modalité déontique – afin d’apporter des nuances à ce verbe encore rarement relevées.

4.4.4. Les valeurs modales de *falloir*

Les études sur les valeurs modales de *falloir* sont peu nombreuses, ce qui rejoint notre réflexion de ce début de section. Ceci peut être dû au fait qu’il est plus fréquemment étudié pour sa catégorisation en tant que verbe impersonnel – ce que nous mentionnions en 4.4. – ou alors parce que sa classification parmi les marqueurs de la valeur modale déontique n’a pas encore été remise en question. En effet, lorsque *falloir* est intégré dans les études sur la modalité, il est systématiquement rangé dans la section des marqueurs de la modalité déontique aux côtés de *devoir*, que ce soit dans des travaux classiques et généraux sur la modalité (Gosselin 2010 ; Le Querler 1996), ou dans les recherches sur la notion d’obligation en général (voir Richard 2005) pour l’analyse des marqueurs d’obligation dans des textes juridiques). Rossari (2020) classe également ce verbe aux côtés de *devoir* dans son modèle triadique en attribuant aux deux verbes une CS de *nécessité*. Cependant, *falloir*, contrairement à *devoir*, ne posséderait pas la propriété d’une lecture rhétorique portant sur le *dire*.

C’est donc cette lecture déontique qui domine dans les réflexions sur *falloir*. Patard (2015) rappelle que pour « *falloir*, qui est un verbe déontique, le sens dominant est très logiquement également déontique : la construction sert à exprimer une obligation morale. » (Patard 2015 : 148). Toutefois, avec les effets de sens de *quasi-certitude* ou de *conjecture* accordés par Valiukienė (2022) – présentés dans la sections précédente –, nous pouvons admettre que *falloir* peut exprimer les valeurs modales épistémiques.

Les valeurs modales descriptives sont plus difficiles à accorder à *falloir* car nous pouvons, dans une certaine mesure, toujours trouver une obligation. Ainsi si nous reprenons ici l’exemple 26 :

35. Pour que le vin soit bon, il *faut* un été sec.

Nous pouvons ici comprendre que c’est une *contrainte du monde* mais également une *nécessité* générale.

Les recherches se consacrant aux valeurs modales de *falloir* sont ainsi relativement limitées. Il existe, cependant, un nombre plus conséquent d’études se focalisant sur les phénomènes de grammaticalisation dont il fait l’objet.

4.4.5. La grammaticalisation ou désémantisation de *falloir*

Les études concernant les phénomènes de grammaticalisation pour *falloir* sont nombreuses. Nous commencerons par présenter quelques réflexions générales par rapport à *falloir* et son association avec les *verba cogitandi* et *dicendi*⁷⁵, pour ensuite aller plus en détails sur certaines associations spécifiques comme *faut croire* (Rossari 2012) et *faut dire* (Pusch 2007)⁷⁶.

4.4.5.1. Quelques remarques générales

Grâce aux outils de la linguistique de corpus⁷⁷, Valiukienė (2022) relève que *falloir* est souvent associé à des verbes *psychologiques* de l'ordre de la *pensée* ainsi que des verbes de *communication*⁷⁸. Ces catégories de verbes sont d'ailleurs celles propices à donner naissance à des MD lorsqu'ils sont associés de manière significatives avec d'autres verbes, comme nous le soulignons avec l'étude de Dostie (2004) présentée dans la section 4.1.1. Valiukienė (2022) confirme cela dans son étude car elle écrit que lorsque *falloir* est suivi par un verbe de cet ordre-là, il « perd son poids sémantique de base dans des constructions qui sont plus ou moins figées aux niveaux lexico-sémantique et pragmatique » (Valiukienė 2022 : 11-12). Elle n'utilise pas la notion de MD, mais au vu des exemples qu'elle présente, *falloir* semble effectivement devenir un MD selon les critères de Dostie (2004). Il aurait, dans ce type de construction, une valeur évidentielle car il permet de mettre en évidence une inférence, ou, comme le disait Dostie (2004), d'attirer l'attention de l'interlocutrice :

36. Il *faut croire* que la mémoire emprunte des chemins étonnants et qu'elle a besoin de chocs émotionnels pour réveiller ce qu'elle enfouit dans sa besace. (Valiukienė 2022 : 8)

Valiukienė (2022) note également une association importante entre *falloir* et des verbes de perception comme *voir* et *entendre* :

⁷⁵ Selon la dénomination de Valiukienė (2022) qui regroupe notamment les verbes *croire*, *voir*, *entendre* et *dire*.

⁷⁶ Signalons également qu'il existe une étude de la locution *falloir mieux* (Patard 2015) que nous ne présenterons pas en détail dans cette thèse pour deux raisons : 1) elle ne concerne pas l'association privilégiée de *falloir* avec un verbe mais un adverbe et 2) Patard (2015) admet elle-même que cette forme est peu fréquente et serait en fait un amalgame avec la locution *valoir mieux*.

⁷⁷ Nous expliquerons exactement quelle méthodologie l'autrice a employée en 4.4.6.

⁷⁸ C'est l'autrice elle-même qui nomme ces différentes catégories sans se fonder sur une classification en particulier.

37. **Faut** voir ce qu'elles rayonnent, les filles⁷⁹.

38. Bien, mais il **faut** entendre ce gars de droit ricaner à la cafétéria.⁸⁰

Pour les occurrences avec *voir*, « [n]i le sens de *nécessité*, ni celui de *voir* (au sens de percevoir quelqu'un ou quelque chose par les yeux) ne transmettent la valeur de base. Le sujet invite le destinataire à faire attention à qqn ou qqch » (Valiukienė 2022 : 9). Quant à *il faut entendre* « le sujet insiste auprès fr son interlocuteur pour qu'il fasse attention *aux gens ricaner à la cafétéria*. Il ne l'oblige pas de l'écouter ou percevoir par l'ouïe le ricanement » (Valiukienė 2022 : 9). Dans ces cas, il semble ainsi difficile de retrouver le NS de *falloir*. C'est pourquoi nous pensons que nous pouvons ici admettre que *falloir* subit un phénomène de grammaticalisation plus avancé que ce que nous avons mis en évidence pour *pouvoir* et *devoir* car nous pouvons même le considérer comme un MD et les MD sont par essence grammaticalisés.

4.4.5.2. Faut croire

Rossari (2012) se penche sur trois marqueurs qu'elle qualifie d'évidentiels, soit *faut croire*, *on dirait* ainsi que *paraît*. Pour les deux premiers, elle rappelle que leur valeur évidentielle vient du fait qu'ils combinent une indication modale – le verbe *falloir* et le mode du conditionnel pour *on dirait* – avec « des verbes désignant des attitudes propositionnelles ou énonciatives (resp. *croire* et *dire*) » (Rossari 2012 : 65). Par cette association, « [l]e locuteur signale [...] qu'il ne transmet pas la description d'un état de choses qu'il a perçu directement » (Rossari 2012 : 65). La désémantisation n'est ainsi pas seulement visible sémantiquement, mais également syntaxiquement. En effet, ces locutions peuvent se retrouver détachées à droite⁸¹, voire à gauche des énoncés, mais également former un énoncé à elles seules. Ainsi, *faut croire*, *on dirait* ou encore *paraît* sont des réponses acceptables dans l'exemple suivant :

⁷⁹ Nous avons aussi trouvé une mention de la locution *faut voir* dans Abeillé et Koenig (2021a : 155) qui la traitent sous l'angle des verbes impersonnels et des constructions qui permettent une omission du sujet. Les autrices classent ce type de construction parmi les usages informels de *falloir*, plutôt usités à l'oral. A nouveau, ce n'est pas l'aspect sémantique de *falloir* qui les intéresse ici mais les caractéristiques syntaxiques de ce verbe dans un certain type de construction.

⁸⁰ Notons que les exemples 39 et 40 sont tirés de l'oral, ou du moins, d'une retranscription de l'oral. Ceci confirme un fait attesté depuis longtemps, soit que le phénomène de grammaticalisation est d'abord perceptible à l'oral (Narrog & Heine 2011 : 4).

⁸¹ Rossari (2012) utilise cette dénomination et ne mentionne pas de phénomène de clivage ou de dislocation pour ces constructions.

39. - Henry n'est pas là, il est malade ?

- **Faut croire** / on dirait / paraît.

Rossari (2012) traite la locution *faut croire* en parallèle d'autres associations de *falloir* avec des formes épistémiques, soit *savoir* et *imaginer*, ce qui nous rappelle l'observation de Valiukienė (2022) susmentionnée et donc les remarques sur la grammaticalisation et plus précisément les MD de Dostie (2004). Avec *savoir* et *imaginer*, *falloir* a « un emploi en tant que marqueur détaché à droite sans clitiques et un emploi semblable en tant qu'intervention dialogique » (Rossari 2012 : 71), c'est-à-dire que les locutions *faut savoir* et *faut imaginer* peuvent elles aussi former une réponse à part entière comme *faut croire*. Ce détachement à droite – ou cette autonomie de signification – n'est pas possible lorsque *falloir* est associé à des verbes comme *penser* ou *envisager*, issus de la même catégorie sémantique que les précédents, et qui figurent régulièrement en position de complémentation verbale de *falloir*. C'est donc sur ces deux critères – le détachement et l'autonomie sémantique – que Rossari (2012) fait la différence entre des emplois grammaticalisés ou non.

À partir de ces observations syntaxiques et sémantiques de *falloir* avec des verbes *putandi*, l'autrice explique que l'usage de *faut croire* « permet de confirmer [ce qui] est déjà accessible au moment de l'énonciation » (Rossari 2012 : 73) Si nous schématisons ceci, nous pouvons admettre que :

avec *p*, **faut croire** *q*.

et donc :

avec *p*, on est sous la *contrainte* / on a la *nécessité* de croire *q*.⁸²

Afin de corroborer le statut grammaticalisé de *faut croire*, Rossari (2012) propose également d'observer cette locution dans un cadre diachronique. Elle remarque ainsi des occurrences assez fréquentes de *faut le croire* au 18^{ème} siècle. Dans ce type d'usage,

il faut le croire, utilisé en tant qu'acte de langage à part entière, signale la nécessité qu'il y a à croire en l'état de choses mentionné dans l'énoncé précédent et qui demande à être confirmé. Toutefois, les indices sur lesquels cette nécessité se fonde sont passés sous silence. (Rossari 2012 : 76)

⁸² Cette interprétation permet d'expliquer pourquoi dans le modèle triadique présenté quelques années plus tard par Rossari (2020), *falloir* ne porte pas sur le *dire* mais sur la *croyance*, et n'aurait donc pas encore atteint un stade énonciatif contrairement à *devoir*. Il n'y a donc pas une prise en charge de l'énoncé avec *falloir* contrairement à l'usage épistémique de *devoir* présenté par Rossari (2020).

Il est ainsi important de mentionner que cette lecture est possible en partie grâce au NS de la *nécessité* (ou de *contrainte*) de *falloir*, qui peut encore être récupéré malgré un emploi grammaticalisé :

La nécessité inscrite dans le sens modal de *falloir* est un trait persistant – au sens que les théoriciens de la grammaticalisation donnent à ce terme (cf. Marchello-Nizia 2006) – dans *faut croire* marqueur. Il a infléchi la valeur évidentielle de ce dernier en le contraignant à intervenir dans des contextes où l'état de choses auquel il se rapporte est notoirement "évident", d'où la nécessité d'y croire. (Rossari 2012 : 76)

Donc malgré son emploi grammaticalisé, du moins, en apparence, la locution *faut croire* semble encore véhiculer le NS de la *nécessité* dans certains contextes.

4.4.5.3. Faut dire

Pusch (2007) se penche sur la locution *faut dire* et décortique 106 occurrences tirées de corpus de français parlé afin d'analyser les valeurs modales de cette construction. Il décrète ainsi que, dans trois fois plus de cas, la grammaticalisation de *falloir* avec *dire* s'opère lorsque *falloir* véhicule une valeur modale *épistémique*, car la locution renvoie à un « jugement du locuteur et au niveau pragmatique de l'échange communicatif » (Pusch 2007 : 32). Cependant, cette association n'est pas automatiquement grammaticalisée et *faut dire* peut parfois véhiculer une valeur modale déontique :

40. il savait pas il savait pas conduire quoi *faut dire* les choses claires et nettes⁸³

Pour Pusch (2007), cet exemple véhicule une valeur modale déontique et cela se vérifie car il peut être réfuté :

41. - *faut dire* les choses claires et nettes

- Ce n'est pas vrai : il ne faut pas dire les choses claires et nettes

Selon Pusch (2007), un emploi à valeur modale épistémique, et donc grammaticalisé de *faut dire*, ne peut faire l'objet d'une contestation :

⁸³ Tous les exemples présentés dans cette section sont issus directement de l'article de Pusch (2007) ou en sont inspirés. Ces exemples sont des transcriptions d'expressions orales, d'où l'absence éventuelle de ponctuation ou de négation dans ces extraits.

42. - d'ailleurs à Orléans puisqu'on parle beaucoup des loisirs il *faut dire* qu'à Orléans euh c'est assez c'est assez réduit hein il n'y a pas cette activité que l'on rencontre dans les dans les villes de province
- Ce n'est pas vrai :
- ? il ne *faut* pas *dire* qu'à Orléans c'est assez réduit
- il *faut dire* qu'à Orléans ce n'est pas réduit

Pusch (2007) reconnaît cependant que ce test présente certaines limites et que la négation pourrait être justifiée dans ce contexte.

Pour déterminer si la grammaticalisation a eu lieu, Pusch (2007) propose donc une analyse morpho-syntaxique de cette structure. Ainsi, les occurrences de *faut dire* suivies d'une proposition subordonnée complétive semblent être dépourvues de sens, alors que les cas où cette expression est suivie d'un pronom ou d'un nom exprime une modalité *déontique* et sont donc moins figés.

Un autre élément qui pourrait aider à repérer et donc à argumenter en la faveur d'une grammaticalisation de *faut dire* sont les occurrences où le pronom sujet *il* ou encore l'introduit de complétive *que* sont omis à l'oral⁸⁴. Dans ces cas, la forme se désémantise car « la présence obligatoire d'un sujet étant considérée un trait verbal essentiel du français, la perte du pronom sujet peut être interprétée comme un processus de déverbalisation » (Pusch 2007 : 33).

Mentionnons également que Pusch (2007) fait référence à l'étude de Kronning (1988) sur *je dois dire que* que nous avons présentée en 4.3.5.2. Selon Pusch (2007), l'emploi concessif attribué à *je dois dire que* n'est pas transposable à *faut dire* car « dans la vaste majorité des cas, cette concessivité reste latente et difficilement palpable du fait que l'antécédent de la relation concessive (l'action ou l'état de choses dont le résultat ou la conséquence escomptés ne se sont pas produits) n'est pas repérable » (Pusch 2007 : 38).

Ainsi, la grammaticalisation de *faut dire* a lieu seulement dans des contextes tout comme c'était le cas pour la locution *faut croire* évoquée précédemment.

⁸⁴ Notons que cette omission n'est peut-être pas uniquement due à une grammaticalisation de la forme *faut dire* en particulier mais un fait récurrent pour les verbes impersonnels : « Ces occurrences reflèteraient plutôt une tendance plus générale d'omission du pronom non-référentiel *il* devant les verbes impersonnels, où ce pronom n'est qu'un élément de remplissage d'une position syntaxique non-actancielle imposé par le caractère toujours plus contraignant du sujet dans l'histoire du français » (Pusch 2007 : 34).

4.4.6. Les méthodologies employées pour appréhender les propriétés sémantiques de *falloir*

Comme nous le mentionnions dans l'introduction de cette partie, *falloir* a été moins souvent analysé que *pouvoir* et *devoir*. On peut même dire que

[d]ans la littérature consacrée aux auxiliaires modaux du français, le verbe *falloir* fait un peu figure de parent pauvre. Si on le compare aux verbes modaux *stricto sensu* (*pouvoir*, *devoir*), on trouve seulement quelques indications concernant son statut grammatical [...]. Les études portant sur sa sémantique sont encore plus rares [...]. (Valiukienė 2022 : 2)

Nous avons d'ailleurs mis en avant qu'il est souvent appréhendu à travers les locutions qu'il forme et pour ses particularités syntaxiques de verbe impersonnel et non pas dans le cadre des études sur les verbes modaux.

Dans cette partie, nous allons nous focaliser sur la façon dont il a été appréhendu par les chercheuses lorsqu'elles ont voulu faire émerger son NS et ses effets de sens. La méthode de la traduction est très appréciée (4.4.6.1.) et les outils de la linguistique de corpus ont aussi permis aux spécialistes de faire émerger quelques régularités sémantiques pour *falloir* (4.4.6.2.).

4.4.6.1. La traduction

Les outils de la traduction sont des moyens privilégiés pour faire ressortir les caractéristiques sémantiques de *falloir*. Ceci est dû à sa valence particulière de verbe impersonnel qui ne trouve pas toujours son équivalent dans les autres langues⁸⁵. Ainsi, les chercheuses tentent de lister par quels mots ou groupes de mots *falloir* est traduit afin d'en déduire ses caractéristiques sémantiques.

Pakpahan et al. (2019) ont adopté cette approche dans leur étude de l'indonésien, une langue dépourvue de verbe impersonnel exprimant la *nécessité*. Ils ont également examiné la manière dont l'équivalent indonésien le plus proche de *falloir*, à savoir *harus*, a été traduit en français, afin de déterminer si cette correspondance est établie ou non. A partir de textes traduits de

⁸⁵ A ce sujet, se référer à Bassnett-McGuire (1980), puis Bassnett (2011), qui explique que la traduction peut, certes, amener une perte au niveau de la compréhension mais dans d'autres cas également un gain à ce niveau. En effet, dues aux différences sémantiques, syntaxiques mais aussi culturelles, la traduction peut apporter dans certains cas des précisions de contenu selon le texte traduit. Dans la Bible, par exemple, la version finlandaise spécifie le type de neige décrit, la version arabe détaille les comportements des chameaux, tandis que la version française précise la variété de pain consommée (Bassnett-McGuire 1980 : 30-31). Si la syntaxe entre deux langues n'est pas la même, la traductrice devra également adopter des stratégies pour préciser le contenu de la langue source (Bassnett 2011 : 120-121). C'est dans cette configuration que se situe *falloir* : étant donné que toutes les langues ne disposent pas forcément d'une forme impersonnelle pour exprimer la nécessité, l'étude de sa traduction permet de préciser ses caractéristiques sémantiques dans la langue source, en l'occurrence le français.

nouvelles, Pakpahan et al. (2019) ont dénombré six effets de sens à *falloir* : la *nécessité*, l'*interdiction*, le *souhait*, le *désir*, l'*ordre* et le *conseil*⁸⁶, effets de sens qu'elles ont donc inventorié à l'aide des traductions en indonésien des occurrences de *falloir* en français.

Valiukienė (2022) a adopté une méthode similaire sur un corpus de textes littéraires français traduits vers le lituanien. Cela lui a justement permis de conclure que

tout en gardant sa valeur canonique de la nécessité, [*falloir*] peut exprimer la conjecture et, en tant que composante dans les unités avec les “*verba cogitandi et dicendi*” perdre sa valeur de base et devenir un marqueur pragmatique de connexion et d'argumentation. (Valiukienė 2022 : 1)

Les recours à la traduction peuvent ainsi être éclairants quant aux divers effets de sens qu'un mot peut revêtir selon les cotextes.

4.4.6.2. Les outils de la linguistique de corpus

Plusieurs des études que nous avons abordées tout au long de ce chapitre sur *falloir* ont employé des méthodes de la linguistique de corpus. Parmi celles-ci, nous pouvons rappeler l'étude de Valiukienė (2022) qui a eu recours à un corpus parallèle de textes littéraires traduits du français vers le lituanien pour identifier les diverses nuances sémantiques de *falloir*. Cependant, ces corpus sont principalement exploités pour trouver des équivalents du verbe dans une autre langue afin d'en éclairer les effets de sens et valeurs modales. Comme nous l'expliquions pour *devoir* (4.3.9), Valiukienė (2022) utilise ainsi les outils de la linguistique de corpus de manière qualitative et non pas quantitative⁸⁷.

Patard (2015) recourt également à des corpus pour recueillir des données authentiques sur l'utilisation de *falloir* et plus précisément pour la locution *falloir mieux*, qu'elle analyse dans une perspective diachronique. Contrairement à Valiukienė (2022), Patard (2015) propose quelques calculs basés sur les fréquences relatives des formes – notamment la représentativité de ces occurrences à travers le temps – ou encore la proportionnalité des valeurs modales véhiculées par ces formes. Certes, ces calculs restent relativement simples par rapport à la quantité d'outils qui existent dans la branche de la linguistique de corpus mais ils ont le mérite de représenter de manière quantitative des résultats qui n'auraient pas été visibles sans ce type

⁸⁶ Pour l'effet de sens de *conseil*, il semble y avoir confusion pour les autrices entre *falloir* et *valoir* dans l'expression *vaut mieux*.

⁸⁷ Nous développerons cette dichotomie dans la section méthodologique à venir (chapitre 6). Toutefois, il convient déjà de souligner que, selon notre conception, cette discipline ne se limite pas à l'identification d'occurrences et de traductions authentiques dans les corpus, mais englobe également des analyses statistiques permettant d'appréhender quantitativement les distinctions ou similitudes fondamentales entre les verbes modaux étudiés dans cette thèse.

de manipulation. Cette méthode se rapproche ainsi de celle de Kronning (1996) qui a également quantifié les divers usages de *devoir* dans un corpus (4.3.9.).

4.4.7. Conclusions sur *falloir*

Même si *falloir* intéresse moins souvent les chercheuses en tant que verbe modal, celui-ci a tout de même bénéficié de quelques études éclairantes sur ses caractéristiques sémantiques. Il est ainsi fréquent d'associer *falloir* à *devoir* dans le cadre des modalités déontiques. Cependant, comme cela a été soulevé précédemment, *falloir* semble différer de ce dernier, et la nature impersonnelle de ce verbe pourrait jouer un rôle significatif à cet égard. En effet, selon Caron et Caron-Pargue (2003), le sens de *falloir* serait plus lié à la notion de *contrainte* qu'à celle de *nécessité*, mettant ainsi l'accent sur une influence extérieure sur l'impératif véhiculé plutôt que sur l'agent, comme cela peut être le cas pour *devoir*. Il convient cependant de noter que cette distinction semble avoir peu d'incidence sur l'interprétation des valeurs modales et des effets de sens de *falloir*, qui peuvent être associées aussi bien à une *contrainte* qu'à une *nécessité*. Ainsi, certains effets de sens, tels que la *nécessité*, l'*obligation* ou encore la *conjecture*, sont rattachables tant à *falloir* qu'à *devoir*.

Concernant les valeurs modales que *falloir* peut transmettre, c'est souvent la modalité déontique qui lui est attribué. Toutefois, au vu des effets de sens qui lui sont accordés – comme l'*hypothèse* ou encore la *quasi-certitude* – nous pouvons aussi admettre des emplois à valeur modale épistémique mais difficilement descriptive.

D'ailleurs, c'est à partir de la valeur modale épistémique que se construisent les locutions figées dont fait partie *falloir* et qui a intéressé les chercheuses. Ces locutions sont principalement des constructions dans lesquelles *falloir* est suivi par un verbe de *communication*, de *pensée* ou encore de *perception* qui grammaticalisées.

Notons d'ailleurs que c'est seulement dans le cadre du figement de ces formes que les préférences cotextuelles de *falloir* ont été mentionnées dans les études existantes à son sujet, ce qui n'était pas le cas pour *pouvoir* et *devoir*. Les complémentations verbales de ces modaux permettaient de retrouver les effets de sens ou leurs valeurs modales et non pas de seulement récupérer les formes désémantisées.

Pour finir, mentionnons aussi que *falloir* a également été souvent étudié en faisant recours à la traduction. Il a aussi bénéficié de l'émergence des outils de la linguistique de corpus pour être analysé. Cependant, comme pour *devoir*, ces outils ont surtout permis aux chercheuses

d'extraire des exemples authentiques alors qu'il existe de nombreux autres usages possibles avec la linguistique de corpus.

4.5. *Vouloir*

Vouloir est le dernier verbe modal que nous considérons dans notre étude. Si *vouloir* – et la notion de volonté en général – sont étudiés dans de nombreux autres domaines que la linguistique, notamment la philosophie ou encore la psychologie (Bogacki 2017 : 48), il n'est que très rarement envisagé sous l'angle de la modalité.

De la même manière que pour *pouvoir*, *devoir* et *falloir*, nous allons nous focaliser sur les réflexions sémantiques qui entourent ce verbe. Pour cela, nous commencerons par expliquer pourquoi nous pouvons le considérer comme un verbe modal malgré les réticences de certaines chercheuses à ce propos (4.5.1 et 4.5.2). Nous nous focaliserons ensuite sur les valeurs modales qu'il peut transmettre (4.5.3.), ses divers effets de sens (4.5.4.) et son NS (4.5.5). *Vouloir*, tout comme *falloir*, a également fait l'objet de beaucoup de recherches concernant sa désémantisation, voire sa grammaticalisation, et nous exposerons ainsi les différentes locutions dans lesquelles il perd son NS (4.5.6.). Nous finirons notre revue de la littérature sur *vouloir* avec les méthodologies qui ont été employées pour en faire ressortir ses particularités sémantiques (4.5.7.).

4.5.1. Pourquoi *vouloir* ne serait pas un verbe modal

Le verbe *vouloir* – ou plus généralement le concept de volition – n'est pas toujours pris en compte dans les études sur les verbes modaux, notamment à cause de la vision étroite de la modalité, encore souvent adoptée dans les analyses⁸⁸. D'ailleurs, dans la GGF, Abeillé et Koenig (2021a : 167) classent *vouloir* dans les verbes de *désir* ou de *sentiment*, alors que *pouvoir*, *devoir* et *falloir* sont qualifiés de modaux. Sctrick (2024) a d'ailleurs soulevé cette problématique dans son article sur la modalité dans l'encyclopédie *Universalis* en écrivant que « c'est une question non résolue de savoir si “vouloir” [...] fait partie des modaux ou s'il s'agit d'un verbe à statut plein comme “désirer” ou “souhaiter” » (Sctrick 2024).

En français, il est vrai que *vouloir* n'occupe « qu'une place marginale dans le système modal » (Larrea 2004 : 740). Tout comme *falloir*, il possède une syntaxe qui le distingue des autres

⁸⁸ Voir la distinction entre vision étroite et vision large de la modalité dans l'introduction de ce travail (1).

verbes modaux. En effet, *vouloir* peut avoir un complément de type syntagme nominal (SN)⁸⁹ ce qui n'est le cas ni de *pouvoir* ni de *devoir*. Cette caractéristique est toutefois commune à *falloir* (Riegel et al. 2008 : 324)⁹⁰. Il serait ainsi peu judicieux de discréditer *vouloir* mais de considérer *falloir* sur la base de propriétés syntaxiques. Mentionnons également que dans les constructions avec un SN, le verbe à l'infinitif est en fait simplement sous-entendu⁹¹. Lorsqu'une locutrice énonce :

43. Je **veux** une glace.

Elle sous-entend :

44. Je **veux** manger/acheter une glace.

La même réflexion vaut pour *falloir* :

45. Il **faut** un passeport pour traverser cette frontière

Exemple qui pourrait être précisé avec un verbe à l'infinitif :

46. Il **faut** posséder un passeport pour traverser la frontière.

Une autre particularité de *vouloir* face aux autres modaux serait sa capacité à introduire une complétive en *que*. Cependant, ceci ne le différencie pas tant que cela des autres verbes de notre sélection puisque c'est également le cas pour *falloir* et, dans une moindre mesure, *pouvoir*. En effet, dans sa forme pronominale et impersonnelle, nous pouvons admettre un usage de la sorte :

⁸⁹ Dell'Oro (accepté-b) différencie d'ailleurs les formes pré-modales et modales en partant de l'équivalent latin de *vouloir*, *uolo*, en signalant que les premières concernent les constructions suivies d'un SN, et qui exprimerait donc un désir (*je veux cette voiture* = *j'ai le désir de posséder cette voiture*) et les secondes lorsque le verbe est suivi d'un verbe à l'infinitif (*je veux acheter cette voiture* = *j'ai l'intention d'acheter cette voiture*). L'intention, qui peut être considéré comme le NS de *vouloir* (4.5.5.), justifie ainsi son intégration dans le paradigme des verbes modaux.

⁹⁰ Dans une certaine mesure, nous pouvons également admettre des constructions de *pouvoir* suivies d'un SN. Celles-ci sont plus rares que pour *falloir* et *vouloir* et ne semblent être apparues que récemment dans la langue orale. Par notre expérience personnelle, nous avons pu entendre *Je **peux** le dernier bout de pain ?* construction dans laquelle le verbe à l'infinitif *prendre* est déduit du contexte. Nous n'avons pas trouvé d'étude ou de remarque à ce propos dans les études parcourues pour ce travail.

⁹¹ Nous empruntons le terme *sous-entendu* aux autrices de la GGF qui, dans plusieurs articles, utilisent cette terminologie pour parler notamment du complément verbal qui est sous-entendu lorsqu'il est omis (Koenig & Abeillé 2021).

47. Il se **peut** que Barbie gagne l'Oscar du meilleur film cette année.

Par rapport aux autres verbes modaux, *vouloir* est également plus souvent considéré comme un verbe sémantiquement « plein » et aurait ainsi l'habileté de sélectionner son sujet, ce qui ne serait pas le cas des autres modaux lorsqu'ils sont considérés comme des *semi-auxiliaires*. Cette différence rejoint la classification d'Abeillé et Koenig (2021a : 169), déjà mentionnée dans l'introduction de ce chapitre, catégorisant les verbes demandant un complément sous forme de verbe à l'infinitif dans deux groupes, les verbes à *partage* et à *héritage*. *Vouloir* fait partie de la première des catégories puisqu'il partage le sujet grammatical avec le verbe à l'infinitif qu'il modalise et « [a]insi, avec un verbe comme *vouloir*, le sujet [...] est interprété comme un agent (celui qui veut) » (Abeillé & Koenig 2021a : 169). Les autres verbes de notre étude, soit *pouvoir* et *devoir* sont des verbes à *héritage*, tandis que la nature impersonnelle de *falloir* ne le fait même pas entrer dans de telles considérations. Cependant, Abeillé et Koenig (2021a) précisent également que « [d]e façon générale, dès qu'un verbe à partage peut être employé avec un sujet non humain, non agentif, et dans un emploi statique, il tend à avoir des propriétés d'un verbe à héritage » (Abeillé & Koenig 2021a : 171) et nous verrons dans les sections suivantes que la *volonté* exprimée par *vouloir* ne se limite pas aux êtres humains. En réalité, *vouloir* est ainsi beaucoup plus similaire aux autres verbes modaux que ne le pensent les autrices qui cherchent à le différencier par ses caractéristiques syntaxiques.

Concernant l'aspect impersonnel des verbes modaux, Blanche-Benveniste (2002 : 85) souligne que le verbe *vouloir* se distingue des autres par son incapacité à être utilisé sous une forme impersonnelle. Il semble que Blanche-Benveniste (2002 : 85) n'avait pas connaissance des emplois régionaux de *vouloir* dans son étude :

48. Il se **peut** qu'il pleuve ce soir.

49. Il **doit** pleuvoir ce soir.

50. Il **faut** qu'il pleuve ce soir.

51. Il **veut** pleuvoir ce soir.

L'usage impersonnel de *vouloir* en 51 est attesté seulement dans certaines régions francophones – la Suisse francophone et la Franche-Comté – et dans des contextes particuliers. Gadet (2021) le qualifie de *futur périphrastique*. Pourtant, il semble y régner une certaine forme d'incertitude dans cet usage qui pourrait donc s'associer à la modalité épistémique, en particulier dans les exemples cités par Gadet (2021) :

52. Elle ne pourra pas étendre aujourd'hui parce que je sens que ça **veut** pleuvoir.

53. Descends de là, tu **veux** tomber.

Exemples de Gadet (2021 : 1273).

Ni la pluie, ni la chute ne semblent être des événements certains d'arriver, seulement probables et peuvent donc être classés comme véhiculant une valeur modale épistémique.

Pour toutes ces raisons mentionnées ci-dessus, il est donc légitime de considérer *vouloir* comme un verbe modal.

4.5.2. Pourquoi *vouloir* est un verbe modal

Vouloir est considéré comme un verbe modal pour les chercheuses qui adoptent une vision large de la modalité⁹². Dans cette vision, toute « expression de l'attitude du locuteur par rapport au contenu propositionnel de son énoncé » (Le Querler 2004 : 646) est considérée comme un marqueur de la modalité. Ainsi, la notion de *volonté* permet « d'exprimer un "état" de la personne, une disposition intérieure, [...] préalable au procès » (Leeman-Bouix 2002 : 127).

Dans une vision étroite de la modalité, certaines autrices arrivent également à intégrer *vouloir* dans leur paradigme. C'est notamment le cas de Larreya (2004) qui explique que la volonté joue un rôle essentiel dans l'articulation des notions de *possibilité* et de *nécessité* exprimées par *pouvoir* et *devoir* :

Même si la notion VOULOIR ne fait pas partie de la modalité au sens le plus strict du terme, il est difficile de concevoir les notions POUVOIR ou DEVOIR sans la faire intervenir, du moins au niveau de l'implicite. Par exemple, *Est-ce que je peux t'emprunter ton mobile ? May I borrow your mobile ?* équivaut, pour au moins une partie de son sens, à *Est-ce que tu veux bien que j'emprunte ton mobile ?* De même, *Il faut que tu m'écoutes ! You must listen to me!* équivaut partiellement à *Je veux que tu*

⁹² Cela semble d'ailleurs être plus souvent le cas dans les études anglophones. Ainsi Van der Auwera et Plungian (1998) déclarent explicitement qu'ils prennent en compte la modalité volitive et l'évidentialité dans leur paradigme des modalités. Harner et Khemlani (2020), quant à elles, expliquent que les verbes de volonté traduisent l'attitude de la locutrice et c'est pour cette raison qu'on peut les intégrer dans des études sur la modalité. Nuyts (2016) exprime d'ailleurs son étonnement face au fait que la volition n'est pas systématiquement prise en compte dans les études sur la modalité : « La raison pour laquelle cette catégorie n'est pas systématiquement intégrée dans la discussion des notions modales reste floue. Peut-être que tout le monde ne la considère pas comme une catégorie modale. Ou peut-être a-t-elle simplement échappé à une attention plus large jusqu'à présent, possiblement parce que ce sens n'est pas très marqué (bien qu'il ne soit pas absent – voir Kratzer 1978 ; Nuyts et al. 2010) dans les auxiliaires modaux en anglais ou dans d'autres langues d'Europe occidentale (la sémantique des auxiliaires modaux en particulier a largement dominé la discussion sur la sémantique de la modalité dans son ensemble). Cependant, il existe de nombreuses expressions lexicales (verbaux, adverbiaux, adjectivaux) avec ce type de sens dans ces langues » (Nuyts 2016 : 39). Nuyts (2016) rappelle ainsi que la considération du concept de volition dans l'étude de la modalité dépend des langues. Rappelons qu'en allemand, notamment, le verbe *wollen*, équivalent à *vouloir*, est systématiquement pris en compte dans les études sur les verbes modaux (Hütsch 2020 : 51).

m'écoutes. Il n'est donc pas surprenant que, dans d'assez nombreuses langues (par exemple en allemand), le système modal inclue un signifiant de VOULOIR. (Larrea 2004 : 739-740)

En cela, il rejoint également la vision de Narrog (2012a : 49) qui intègre la notion de « volitivité » dans ses descriptions des catégories modales. Selon lui, les valeurs modales exprimés par les marqueurs modaux sont plus ou moins empreintes de volitivité. Dans cette optique, les modalités déontiques sont davantage marquées par la volitivité que les modalités épistémiques ou descriptives.

Un des arguments que nous avons également repérés en faveur de l'intégration de *vouloir* dans le paradigme des verbes modaux est la théorie des mondes possibles de Kratzer (1981) que Martin (2005 : 11) rappelle dans son étude :

Pierre veut revenir : la tradition range un tel énoncé parmi les énoncés modaux. Le fait est que l'auxiliaire *vouloir* situe la proposition *Pierre revient* dans les mondes possibles : cette proposition n'est pas une proposition vraie dans le monde de ce qui est : intuitivement un tel énoncé relève de la modalité. (Martin 2005 : 11)

Ceci le rapproche ainsi des autres verbes mentionnés dans notre recherche qui, comme expliqué en 4.2.1., s'intègre également dans cette théorie.

En résumé, bien que *vouloir* soit souvent exclu des études sur les verbes modaux – notamment par les chercheuses qui adoptent une vision étroite de la modalité – les arguments présentés dans ces deux sections démontrent que *vouloir* partage de nombreuses caractéristiques avec les verbes modaux tels que *pouvoir* et *devoir*. Non seulement *vouloir* participe à l'expression de la volonté, une attitude de la locutrice par rapport au contenu propositionnel de son énoncé, mais il s'intègre également dans la théorie des mondes possibles, tout comme les autres modaux étudiés.

4.5.3. Les valeurs modales de *vouloir*

Généralement, *vouloir* est considéré comme un marqueur des modalités bouliques. Celles-ci « servent à exprimer des désirs, des volontés, des souhaits » (Gosselin 2010 : 351). Le Querler (1996) classe la notion de *volonté* dans les modalités subjectives aux côtés de la notion d'*appréciation*. Nuyts (2016) utilise le terme « préférence » lorsqu'il définit les modalités bouliques.

Comme nous l'avons relevé ci-dessus, *vouloir* a aussi la capacité de transmettre d'autres valeurs modales notamment la modalité déontique car il permet, par exemple, d'exprimer un ordre et

est ainsi étudié dans le cadre des actes illocutoires directifs (Kissine 2007 : 162), acte souvent atténué à l'aide de l'imparfait :

L'imparfait peut aussi exprimer une demande polie dans une proposition principale, avec un verbe de volonté ou de mouvement à la première personne : *Je voulais/venais vous demander un service.*

La demande est doublement atténuée : on ne la présente *pas* directement, mais sous la forme d'une de ses prémisses (volonté ou mouvement vers le destinataire), elle-même décalée dans le passé.

(Riegel et al. 2008 : 310)

Il existe également des tournures à l'oral comme *voulez-vous bien vous taire* « qui permet [...] au locuteur d'exprimer un ordre. *Vouloir* renvoie à tout ce que j'ai déjà dit au sujet des synonymes exiger, ordonner, etc. » (Domínguez 2004 : 906). Dans ces cas, ce n'est pas *vouloir* en tant que tel qui véhicule l'*ordre* mais ce qui est inféré à partir de l'énoncé. Ce genre d'interprétation ramène à la théorie des actes de langage puisque Searle (1979) écrivait déjà que « [l]e *désir* ou la *volonté* regrouperont les requêtes, les ordres, les commandements, les demandes, les prières, les plaidoyers, les supplications et les adjurations » (Searle 1979 : 43) permettant alors de corroborer les interprétations à valeur modale déontique pour *vouloir*. Soulignons que dans ces cas, ce sont à nouveau les procédés pragmatiques, notamment la théorie de la politesse de Brown et Levinson (1987), qui peuvent être convoqués pour interpréter *vouloir*. Cette théorie explique, entre autres, que l'expression des désirs doit justement se faire en maintenant sa face et celle de son interlocutrice : « Étant donné que la face consiste en un ensemble de désirs qui ne peuvent être satisfaits que par les actions (y compris l'expression de ces désirs) des autres, il sera généralement dans l'intérêt mutuel de deux participants à la conversation de préserver la face de chacun » (Brown & Levinson 1987 : 60). Parmi les actes qui peuvent être néfastes pour la préservation de la face figurent notamment les ordres et les requêtes (Brown & Levinson 1987 : 66). Il n'est donc pas anodin que des procédés linguistiques soient utilisés pour les adoucir, notamment l'utilisation de *vouloir* pour exprimer une demande de faire, comme nous venons de l'observer ci-dessus. L'expression de la valeur modale déontique de *vouloir* ne peut être appréhendée que par le biais des théories pragmatiques et doit donc être inférée à partir du contexte. Cette valeur modale ne fait pas intrinsèquement partie du NS de *vouloir*. Il est ainsi important de différencier les valeurs modales et effets de sens directement percevables à partir du NS et ceux qui apparaissent à l'aide de certains contextes particuliers, compréhensibles uniquement à l'aide d'inférences.

Dans une moindre mesure, *vouloir* permet aussi de véhiculer une modalité épistémique lorsqu'il est utilisé en tant que marqueur de futur incertain⁹³.

Ces différentes valeurs modales peuvent ensuite prendre plusieurs effets de sens que nous allons détailler ci-dessous.

4.5.4. Les effets de sens de *vouloir*

La polysémie de *vouloir* est attestée depuis l'antiquité (Goldschmitt & Landvogt 2008 : 255). Riegel et al. (2008) l'ont relevé dans leur grammaire présentant le verbe *vouloir* avec les effets de sens suivants :

vouloir exprime d'abord la volition (*elle veut partir*), et correspond à l'auxiliaire de mode allemand (*Modalverb*) *wollen*. Mais il peut, comme *pouvoir* et *devoir*, concurrencer *aller* pour exprimer un procès à venir teinté d'une volonté figurée (*ce mur veut tomber*), y compris pour des phénomènes météorologiques : *cette année, il ne veut pas pleuvoir*⁹⁴ : dans l'expression de l'avenir, la volonté et l'imminence sont liées (voir l'anglais *will*, auxiliaire du futur). (Riegel et al. 2008 : 254)

C'est l'effet de sens de *volition* qui est le plus communément accordé à *vouloir* (Blumenthal 2012 ; Bogacki 2017 ; Deboës 2002 ; Domínguez 2004 ; Goldschmitt & Landvogt 2008 ; Hütsch 2020 ; Riegel et al. 2008). Selon Bogacki (2017), le verbe *vouloir* est même considéré comme le verbe prototypique de cette notion, en contraste avec d'autres verbes susceptibles d'exprimer un effet de sens similaire. Cette composante *volitive* de *vouloir* n'est pas nouvelle et remonte à son équivalent latin :

Étymologiquement, le verbe *vouloir* remonte à *VOLERE* en latin vulgaire provenant de la forme *VELLE* du latin classique. Dans l'histoire de la langue, le noyau sémantique de ce verbe - la volition - est resté relativement stable. Ainsi, la forte majorité des occurrences a une signification volitive, en latin et en français [...] (Goldschmitt & Landvogt 2008 : 255)

En outre, l'effet de sens de *volition* dépend de beaucoup de paramètres, et dans certains de ses emplois, *vouloir* peut ainsi être rapproché des verbes désidératifs comme *souhaiter* et *désirer* :

Il est clair que l'expression de la volition en français revêt une multitude de formes, que cela concerne les synonymes de *vouloir*. À tout cela il faudrait ajouter les "accidents" du verbe : le temps, la modalité et l'aspect. Ainsi pour exprimer le désir le plus rudimentaire *souhaiter* et *désirer* sont les synonymes de *vouloir* conjugué au conditionnel ou au futur antérieur. Cette grande variété –

⁹³ Voir les exemples 52 et 53 et les réflexions que nous avons proposées sur la base de Gadet (2021) en 4.5.1.

⁹⁴ Ceci n'est pas sans rappeler les emplois que mentionne Gadet (2021) dans les exemples 52 et 53.

autant des signifiants et des signifiés –, est la principale caractéristique de l’expression de la volition en français. (Domínguez 2004 : 909)

Les verbes pouvant exprimer la volonté ont fait l’objet de beaucoup de recherches (Blumenthal 2012 : Bogacki 2017 : Harner 2016 : Harner & Khemlani 2020) mais celles-ci ne concernent pas l’intégration de *vouloir* dans un paradigme des verbes modaux ni forcément l’étude de ses effets de sens en tant que tels. Pourtant, sur le plan syntaxique, Riegel et al. (2008) relèvent que les verbes de volonté comme

désirer, espérer, préférer, souhaiter et *vouloir* sont directement suivis de l’infinitif (la langue classique insérait *de* devant l’infinitif) : *Elle désire vous connaître – J’espérais trouver un hôtel un peu plus confortable* (Gide). Ces verbes ont une étroite affinité avec les auxiliaires modaux comme *devoir* et *pouvoir*. (Riegel et al. 2008 : 337).

Concernant les autres effets de sens de *vouloir*, Desclés (2003) détermine que *vouloir* peut exprimer un *souhait*, une *volonté*, ou encore une *demande d’autorisation* (Desclés 2003 : 57-58). La *volonté* pourrait ainsi s’associer à un *ordre* (54) et la *demande d’autorisation* à une permission (55) :

54. Les parents **veulent** que leurs enfants soient à la maison pour le souper.

55. Loanne signale à ses collègues qu’elle **veut** prendre le train de 16h afin que l’horaire de la réunion soit adapté.

C’est à nouveau par inférence que ces effets de sens peuvent être déduits. L’exemple 54 montre bien un ordre dans lequel les parents exigent de leurs enfants qu’elles soient rentrées à l’heure pour le repas. Pour 55, c’est également le co(n)texte qui nous montre que la volonté de Loanne l’amène à demander une autorisation à partir avant ses collègues. La notion de con(n)texte n’est d’ailleurs pas considérée chez Desclés (2003) lorsqu’il attribue à *vouloir* la propriété d’exprimer ce type d’effets de sens qui semble ainsi « inné ».

Deboës (2002) s’est également intéressé aux différents effets de sens de *vouloir*. Il en a détecté trois : « On peut distinguer trois manifestations sémantiques de la volition dans le *Dictionnaire*⁹⁵ : le *vouloir qqch* (le procès expérientiel), *dire de vouloir* (l’acte illocutoire), *le faire qu’on voulait* (l’intentionnalité) » (Deboës 2002 : 309). Le *vouloir qqch* « exprime la perception d’un manque ou d’une absence à laquelle le sujet répond par la projection d’une

⁹⁵ Chez Deboës (2002), le *Dictionnaire* fait référence au « *Dictionnaire sémantique et syntaxique des verbes français*, sous la direction de Halina Lewicka et Krzysztof Bogacki, édition de 1983 à Varsovie. » (Deboës 2002 : 14)

situation méliorative » (Deboës 2002 : 310) et serait ainsi l'expression du *désir*. Le *dire de vouloir* correspond à « une fonction pragmatique illocutoire où x a les capacités d'imposer sa volonté » (Deboës 2002 : 312) qui serait ainsi l'expression d'un *ordre*. Enfin, l'*intentionnalité* est fortement liée à l'effet de sens de *désir* : en effet, un désir peut pousser à l'action (Deboës 2002 : 315).

Cette intention peut aussi se traduire par l'effet de sens de *futur* pour *vouloir* comme l'ont rappelé Riegel et al. (2008), cités au début de cette section. Cependant, « de nos jours, [cette utilisation de *vouloir*] se trouve encore [seulement] dans certains dialectes de l'est de la France et en Suisse » (Goldschmitt & Landvogt 2008 : 255)⁹⁶.

Les terminologies de ces effets de sens ne sont pas toujours les mêmes. Ainsi, Goldschmitt et Landvogt (2008) parlent plutôt d'effets de sens de *désir* et d'*intention* pour les effets de sens *volitif* et de *futur* tels que présentés ci-dessus. Elles précisent d'ailleurs que ces deux effets de sens « ne se distinguent pas toujours nettement » (Goldschmitt & Landvogt 2008 : 256). Ainsi, comme pour *pouvoir*, *devoir* et *falloir*, il peut être compliqué de distinguer les effets de sens d'un verbe modal dans un co(n)texte parfois retreint.

Nous avons ainsi pour *vouloir*, non seulement des effets de sens lié à un *désir* ou à une *volonté*, mais aussi un effet de sens porté sur le *futur* avec notamment la notion d'*intention* qui apparaît chez de nombreuses autrices citées ici. Cette liste des divers effets de sens de *vouloir* nous permet de présenter les réflexions autour du NS de ce verbe.

4.5.5. Le noyau sémantique de *vouloir*

Comme nous venons de le montrer à travers ses divers effets de sens, *vouloir* permet généralement d'exprimer la *volonté* ou, dans certains cas plus rares, le *futur*. C'est pour cela que plusieurs autrices admettent pour *vouloir* le NS d'*intention* qui semble pouvoir faire le lien entre la notion de *volonté* et de *futur*. C'est le cas notamment pour Blumenthal (2012) qui écrit que

⁹⁶ Heine (2002), dans son étude sur la grammaticalisation, précise que c'est commun pour les marqueurs volitifs de marquer à un moment donné le futur, et ce dans plusieurs langues : « Il est communément admis qu'un schéma de volition du type [X veut Y] constitue l'une des trois principales sources pour le développement des catégories de temps futur (Bybee, Pagliuca et Perkins 1991 : Bybee, Perkins et Pagliuca 1994). Cependant, ce schéma peut également donner lieu à une fonction grammaticale différente : celle de l'aspect proximatif. La fonction de cet aspect est de définir une phase temporelle immédiatement antérieure à la frontière initiale de la situation décrite par le verbe principal. Les aspects proximatifs sont couramment traduits par des expressions telles que "presque", "prochainement", "être sur le point de", "être à deux doigts de", et autres expressions similaires » (Heine 2002 : 90).

[c]ertaines passions, comme le *désir*, l'*espérance* et la *crainte*, nous portent "à regarder l'avenir", ce qui implique un jugement de valeur (la chose considérée est bonne ou mauvaise, ce qui nous incite à un acte de volonté) et un jugement de probabilité (d'"apparence"), puisque l'avenir n'est pas sûr. (Blumenthal 2012 : 34)

Ainsi, une *volonté* ou un *désir* sont forcément tournés vers le *futur* car ils ne sont pas encore réalisés. Nous pouvons ainsi avoir une *intention* de réaliser ces *désirs* et *volontés*⁹⁷.

Cependant, il existe des autrices qui ne mettent pas en lien la notion d'*intention* avec le verbe *vouloir*. Leeman-Bouix (2002), par exemple, mentionne l'effet de sens de *futur* mais parle quand même de verbe de *volition* lorsqu'elle décrit *vouloir*. Selon Goldschmitt et Landvogt (2008), c'est la notion de *volonté* qui décrit le verbe (Goldschmitt & Landvogt 2008 : 255). Bogacki (2017) qualifie également *vouloir* comme étant un des verbes prototypiques de la notion de *volonté*.

En somme, le NS de *vouloir* ne fait pas l'unanimité auprès des chercheuses. Ce manque d'accord peut aussi être dû aux nombreux usages désémanés et grammaticalisés de *vouloir*.

4.5.6. La grammaticalisation de *vouloir*

Vouloir a fait l'objet de nombreuses études qui traitent de sa grammaticalisation. La locution la plus fréquemment analysée est *vouloir dire*. Pour Franckel (2017), cette construction confère à *vouloir* un emploi particulier puisque c'est la seule construction dans laquelle le verbe peut posséder un sujet inanimé mais être employé à l'affirmative. Dans tous les autres cas, les sujets inanimés apparaissent avec une négation : « *Le moteur ne veut pas démarrer ; Le café ne veut pas passer ; Cette tâche ne veut pas partir ; Le sommeil ne veut pas venir* » (Franckel 2017 : 9). Précisons que pour les phénomènes de grammaticalisation, « [p]eut-être que l'un des types de changement de sens les plus courants implique qu'un rôle de participant typiquement réservé aux humains soit ouvert à des participants inanimés » (Heine 2002 : 87).

Ces usages ne sont pas sans rappeler la réflexion de Van de Velde (2010) autour des sujets inanimés que nous avons présentée en 4.2.4. Elle explique qu'en position sujet, les noms d'*instruments* peuvent acquérir les propriétés d'un nom animé : l'*intention* ou la *volonté*, généralement attribuées à un être animé peuvent ainsi s'appliquer à un *objet* lorsqu'il est en position de sujet. C'est pourquoi nous pouvons estimer que, dans ce type d'exemple, il ne s'agit pas d'une perte totale du NS de *vouloir*. En effet, l'*intention* est simplement attribuée à quelque

⁹⁷ Cette interprétation est également possible dans d'autres langues comme l'anglais, par exemple, où Condoravdi et Lauer (2016) déterminent pour *want* un NS d'*intention*. Nuyts (2016) met également au même niveau les notions de *volition* et d'*intention* dans sa classification des valeurs modales.

chose de non-humain. Cette réflexion va dans le sens de la notion d'assouplissement présentée par Recanati (2004) :

Il y a assouplissement lorsqu'une condition d'application intégrée dans le concept littéralement exprimé par un prédicat est contextuellement abandonnée, élargissant ainsi l'application du prédicat. Un exemple est : 'Le distributeur a avalé ma carte de crédit.' Un distributeur automatique ne peut pas vraiment avaler, puisqu'il n'est pas un organisme vivant doté de l'équipement corporel nécessaire pour avaler. En assouplissant les conditions d'application de 'avalé', nous construisons un concept *ad hoc* avec une application plus large. (Recanati 2004 : 26)

Ainsi, nous retrouvons encore la notion d'*intention* dans ces usages, mais par *assouplissement* de ce sens qui amène une dimension métaphorique à *vouloir* lorsqu'il est précédé par un sujet non animé.

Si l'on revient à *vouloir dire* en tant que tel, il est important de souligner que la locution n'est pas désémanisée systématiquement. En effet, lorsque

vouloir ainsi que *dire* sont remplaçables par une série d'autres verbes [comme] :

je veux (tiens à, me fais un devoir de, me dois de, souhaite, etc.) te dire (déclarer, prévenir, faire savoir, informer, etc.) clairement que les choses ne peuvent continuer ainsi : ce que je veux dire aux Français, c'est que la peur n'est pas bonne conseillère : voilà ce que je voulais vous dire...

vouloir est associé à une intention ciblée de l'énonciateur sujet de *dire*. (Franckel 2017 : 12)

Franckel (2017) résume sa réflexion en expliquant que « [v]*ouloir dire* correspond à la reprise et à la reformulation d'un dire en cours : il s'agit de *ce que je veux dire en disant ce que je dis (ou viens de dire)* » (Franckel 2017 : 14). Cette locution peut s'utiliser dans plusieurs cas : 1. *J'veux dire* ponctuant, 2. *Ce que je veux dire par là* qui agit comme « un écart établi à partir d'un dit validé qui fait l'objet d'une reprise marquée par *par là* » (Franckel 2017 : 15), 3. *vouloir dire* comme une reprise de rectification et 4. pour expliciter le sens d'un mot.

Plus tard, Franckel (2018) s'est également intéressé à la locution *je vois ce que tu veux dire* en analysant comment ces verbes se combinent entre eux pour donner un nouveau sens. Dans cette locution, *vouloir dire* garde son sens de *reformulation* mais avec une certaine distance sur le dit en question : « *Vouloir dire* ici ne relève en rien d'une intention de dire, mais d'une interprétation qui se dit dans un rapport d'altérité à ce qui est effectivement dit et qui peut se traduire par une reformulation » (Franckel 2018 : 19).

Goldschmitt et Landvogt (2008) ajoutent l'effet de sens *citationnel*⁹⁸ à *vouloir*. Celui-ci serait dû à un emploi auxiliarisé, et donc grammaticalisé, de *vouloir* :

⁹⁸ Dell'Oro (soumis) démontre dans son étude que cet usage existait déjà pour l'équivalent latin de *vouloir*.

la signification prédominante du verbe *vouloir* n'est plus la volition, mais le jugement du locuteur : il cite l'énoncé d'un premier locuteur. En même temps, le choix du modal *vouloir* indique le doute du rapporteur quant à la vérité des paroles rapportées. (Goldschmitt & Landvagt 2008 : 258)

Au vu de leur description, nous pouvons admettre qu'il s'agit dans ce type d'usage d'une valeur modale épistémique, qui est donc de l'ordre de l'incertitude de la locutrice face à son dire. Parmi les énoncés qu'elles classent comme citationnels, figure l'exemple suivant tiré d'un critique littéraire :

55. Les confessions en ce point, toutes ces pages sur son 'apostasie' nous rapportent la vie que Jean-Jacques ***voulait*** avoir vécue, bien plus que celle qu'il avait vécue en effet.

Qui peut être paraphrasé selon les autrices par

56. Les confessions en ce point, toutes ces pages sur son 'apostasie' nous rapportent ***la vie que Jean-Jacques prétendait avoir vécue***. (Goldschmitt & Landvagt 2008 : 258)

Ici, la lecture volitive de *vouloir* n'est pas possible selon les chercheuses car « le critique du livre fait un commentaire subjectif : il croit que les *Confessions* nous racontent la vie de Rousseau telle que ce dernier prétend l'avoir vécue et non pas la véritable histoire de sa vie » (Goldschmitt & Landvagt 2008 : 258). Cependant, Goldschmitt et Landvagt (2008) admettent dans leur article que l'effet de sens citationnel est rare et qu'une lecture volitive peut être possible dans les énoncés auxquels elles ont accordé cet effet de sens à *vouloir*. En effet, il semble que *projection d'avoir vécu quelque chose* peut se rapporter à l'*intention* de l'avoir vécu.

Domínguez (2004) traite le verbe *vouloir* avec plusieurs composantes, notamment l'adverbe *bien*. Selon lui :

Cette tournure est l'exemple type d'une locution verbale : verbe au sens large, plus adverbe : le tout réductible à un troisième terme : *accepter*, *opiner*, etc. L'adverbe *bien* modalise la volonté de telle sorte que celle-ci n'a plus autant d'emprise sur le Hors-Moi que celle qu'expriment les verbes *exiger*, *ordonner*, etc. La locution exprime la concession : en disant "je veux bien", le sujet de la proposition conforme sa volonté à celle de quelqu'un d'autre. (Domínguez 2004 : 905)

Cet effet de sens de concession est à rapprocher de l'étude de Rossari et Ricci (2021) et du modèle triadique de Rossari (2020). Selon ces recherches, la locution *je veux bien*, s'applique à l'énonciation et donc sur le *dire*. Rossari et Ricci (2021) estiment qu'il y a ainsi trois

conditions pour que la locution *je veux bien* ait une valeur de concession et soit donc grammaticalisée :

la valeur confirmative de *bien* dans la construction non figée idoine (*je veux bien un verre d'eau*), le fait de prendre comme OSA E(P)⁹⁹ quand la construction est figée (*un verre d'eau, je veux bien, mais avec des glaçons*), le fait d'être suivi de *mais* (*un verre d'eau, je veux bien* n'est pas interprété comme concessif mais comme une construction disloquée à gauche répondant à une question comme *Tu veux un verre d'eau ?*). (Rossari & Ricci 2021 : 109)

C'est en cela que nous pourrions expliquer la grammaticalisation de cette locution qui agit comme un MD selon les critères de Dostie (2004).

Dans Franckel (2017), le chercheur dénombre aussi d'autres emplois dans lesquels *vouloir* semble se grammaticaliser :

- 57. Et le monsieur, qu'est-ce qu'il **voulait** ? (contexte : un marchand s'adresse à un client)¹⁰⁰
- 58. Quand tu **veux** ! (contexte : on n'attend plus que toi)
- 59. On peut y aller demain, si tu **veux**.
- 60. C'était un peu son deuxième père, si tu **veux**.

Ces usages ne sont pas détaillés davantage mais simplement mentionnés comme des exemples de grammaticalisation de *vouloir* selon l'auteur. Cependant, nous estimons que le NS d'intention reste perceptible dans tous ces emplois. Le fait que le verbe, ou la locution qu'il compose (comme *quand tu veux* en 58 ou *si tu veux* en 59 et 60), puisse constituer une réponse à elle seule ou soit détaché à droite (selon les critères syntaxiques de Rossari (2012) énumérés pour la locution *faut croire* expliqués en 4.4.5.2.) ne suffit pas, selon nous, à justifier la grammaticalisation de *vouloir* dans ces énoncés.

À ces exemples, s'ajoute aussi un *vouloir* logique selon Larreya (1997) comme dans *pourquoi voudrais-tu que j'aie peur ?* :

Il est important de noter que le "vouloir" exprimé ici par le verbe de la proposition principale n'est nullement un "vouloir" déontique : c'est un "vouloir" en quelque sorte logique (une glose possible serait "Pourquoi ta logique veut-elle que j'aie peur ?"). C'est ce type de "vouloir" que l'on a dans des énoncés comme "Eh oui, qu'est-ce que vous voulez, il faut bien qu'il gagne sa vie !", ou encore (en français familier) dans des dialogues comme "Tu crois que l'OM va gagner ?" "Je veux, oui !". (Larreya 1997 : 162)

⁹⁹ Pour *objet sémantique abstrait* portant sur l'énonciation de la proposition.

¹⁰⁰ Les contextes sont ici précisés par Franckel (2017).

Le verbe *vouloir* paraît ici comme un MD et pourrait, en effet, être ici un exemple d'un phénomène de grammaticalisation.

En somme, *vouloir* est sans doute le modal de notre étude qui fait le plus l'objet de phénomènes de grammaticalisation.

4.5.7. Les méthodologies employées pour appréhender les propriétés sémantiques de *vouloir*

Vouloir a davantage intéressé des chercheuses en tant que verbe de volonté que de verbe modal aux côtés d'autres modaux plus traditionnels comme *pouvoir* et *devoir*. Ainsi, Blumenthal (2012) l'a considéré dans des « expressions modales qui relèvent du jugement de valeur et/ou de la volonté » (Blumenthal 2012 : 32) et l'étudie aux côtés de *aimer que*, *aimer à croire*, *vouloir croire*, *espérer* ou encore *désirer* et *souhaiter*. Dans un corpus littéraire diachronique, tiré de Frantext¹⁰¹, il étudie comment ces formes ont évolué sur la base de leurs fréquences d'apparition au fil du temps. À travers quelques exemples de ce corpus, il propose une analyse sémantique de ces verbes, en particulier pour souligner la différence entre *désirer* et *souhaiter*. Bogacki (2017), quant à lui, a pour but de « dresser la liste de toutes les unités lexicales dont le sémantisme comporte une composante volitive » (Bogacki 2017 : 47) afin de voir quels sont les lexèmes les plus représentatifs de la volition. *Vouloir* est le verbe qui apparaît en tête de liste aux côtés de *désirer*, *souhaiter*, *convoiter* et 54 autres potentiels synonymes qui permettent d'exprimer le concept de volonté. L'auteur utilise aussi des méthodes quantitatives dans son article. Mais, si son étude avait pour base les synonymes du verbe *vouloir*, ce sont les noms tels que *volonté* et *désir* et les adjectifs qui les accompagnent de manière significative qui sont ensuite analysés.

Quant aux autres recherches que nous avons citées dans cette partie consacrée à *vouloir*, elles se font uniquement à l'aide de manipulations d'exemples plus ou moins authentiques. Il est donc important de remarquer le manque de recherches quantitatives visant à affiner nos connaissances sur les caractéristiques de *vouloir*, son intégration dans le domaine de la modalité et par conséquent sa comparaison avec d'autres éléments modaux.

¹⁰¹ La base textuelle Frantext est disponible en ligne : ATILF. (1998-2023). *Base textuelle Frantext*. ATILF-CNRS & Université de Lorraine. <https://www.frantext.fr/>

4.5.8. Conclusions sur *vouloir*

Dans cette partie, nous avons voulu lister les caractéristiques sémantiques du verbe *vouloir*. Comme nous l'avons soulevé, celui-ci n'est pas toujours considéré comme un verbe modal. La syntaxe et la valence de ce verbe le différencie des autres verbes considérés traditionnellement comme modaux. Il semble également être plus riche sémantiquement que d'autres verbes modaux.

Pourtant, nous avons pu démontrer à l'aide des études existantes que ses effets de sens pouvaient se traduire en valeurs modales. Il y a, certes, la modalité boulique, propre à *vouloir*, mais les valeurs modales déontiques et épistémiques peuvent également être véhiculées par ce verbe. Les effets de sens et les valeurs modales n'ont pas été explorés à l'aide du sujet grammatical ou de la complémentation verbale à l'infinitif, contrairement aux trois autres verbes abordés précédemment.

L'effet de sens le plus attesté pour ce verbe est celui de la *volition*. Cela pousse les chercheuses à souvent comparer *vouloir* à d'autres verbes partageant cet effet de sens, voire ce NS. Plusieurs linguistes considèrent que l'*intention* représente le NS de ce verbe alors que d'autres parlent de NS de la *volonté* ou de *volition*. Pour les premières, le NS de l'*intention* englobe plus d'effets de sens que celui de la *volition*. Ce dernier ne permet pas, par exemple, de comprendre l'usage périphrastique ou de futur avec un degré de certitude plus ou moins certains pour *vouloir*.

Vouloir a également fait l'objet de recherches concernant sa grammaticalisation, notamment au sein de la locution *vouloir dire* mais aussi dans d'autres usages moins figés, comme lorsqu'il est précédé par un sujet non-animé.

Soulignons aussi que, comme pour les autres verbes modaux traités ici, *vouloir* n'a pas été analysé avec des outils quantitatifs tirés de la linguistique de corpus.

4.6. Comparaisons entre les verbes modaux

Après avoir exposé les caractéristiques sémantiques individuelles de chacun des verbes modaux examinés dans ce travail, cette section vise à déployer quelques études comparatives. Étant donné que cette thèse cherche également à mettre en évidence les similitudes et les différences entre ces quatre verbes modaux, il est nécessaire d'explorer ce qui a déjà été accompli dans ce domaine et la façon dont ces recherches ont été menées.

4.6.1. Les comparaisons entre *pouvoir* et *devoir*

Les verbes modaux les plus fréquemment comparés sont *pouvoir* et *devoir* (Barbet 2012, 2015 ; Busquets & Denis 2001 ; Hütsch 2020 ; Rossari et al. 2020 ; Sueur 1977, 1979 ; Vetter 2004, 2012). Ils sont considérés comme les verbes dénotant les « deux valeurs modales fondamentales, la possibilité et l'obligation » (Riegel et al. 2008 : 254), et sont ainsi décrits comme les verbes modaux paradigmatiques du français (De Saussure 2014b : 3). Ceci est dû au fait qu'ils représentent les deux extrémités de l'échelle modale classique qui fait le pont entre la notion de *possibilité* et de *nécessité*.

C'est sur le plan épistémique qu'ils sont le plus souvent mis en parallèle. Comme le rappelle Hütsch (2020) :

Les verbes modaux *devoir* et *pouvoir* font l'objet d'analyses comparatives quant à la différence sémantique fine entre leurs emplois épistémiques respectifs [...]. Malgré leurs noyaux sémantiquement distincts, ces deux verbes sont très semblables en ce qui concerne leurs effets de sens épistémiques, sans pour autant être substituables. (Hütsch 2020 : 26)

Lorsqu'ils sont permutables, Desclés (2003 : 64) précise que « *peut* renvoie plutôt à du *possible* tandis que *doit* renvoie plutôt à du *probable* » (Desclés 2003 : 64). Ce serait le cas comme dans l'exemple suivant :

61. Il ***peut/doit*** pleuvoir ce soir.

L'exemple 61 tiré de Riegel et al. (2008 : 254) montre bien que la modalité épistémique peut être traduite par ces deux verbes. Pourtant, les deux emplois ne sont pas complètement synonymiques selon Desclés (2003 : 64) et dans ce genre d'usage, *devoir* « exprime une probabilité plus forte » (Riegel et al. 2008 : 254) alors que « *pouvoir* épistémique, [...] marque une probabilité plus faible, une simple éventualité » (Gosselin 2010 : 440).

Leurs effets de sens respectifs sont également très proches. Ainsi,

[s]i l'on reconnaît généralement que *pouvoir* exprime la *permission* et *devoir* l'*obligation*, pour les autres interprétations de ces deux verbes, on se trouve confronté à un vaste ensemble de notions que les divers auteurs assignent tantôt à l'un, tantôt à l'autre, tantôt à l'un et à l'autre de ces deux verbes : intention, faculté, capacité, possibilité, probabilité, éventualité, approximation, vraisemblance, supposition, doute, hypothèse, incertitude, etc. (Sueur 1979 : 97)

Pour résumer, ces deux verbes semblent être comparés par rapport au degré de certitude qu'ils expriment lorsqu'ils véhiculent une valeur modale épistémique.

4.6.2. Les comparaisons entre *devoir* et *falloir*

Devoir et *falloir* sont les verbes prototypiques de la valeur modale déontique. Riegel et al. (2008) ont d'ailleurs noté que « *falloir* est à rapprocher, pour le sens, de *devoir* [...] mais il s'insère dans des structures différentes » (Riegel et al. 2008 : 285). Ce serait ainsi principalement relativement à leurs propriétés syntaxiques que les deux verbes se différencient. Pour explorer ces nuances, les chercheuses ont examiné les équivalents de ces verbes dans d'autres langues, comme le souligne Valiukienė (2016) avec sa liste conséquente d'études contrastives portant sur ce sujet :

Les deux principaux verbes de la modalité de nécessité en français, à savoir *devoir* et *falloir*, sont déjà analysés et comparés avec ceux dans d'autres langues sous différents aspects. *Devoir* et *falloir* sont comparés avec les verbes modaux anglais *must*, *have to* (Rivière 1984; Tasmowski & Dendale 1998; Larreya 2004), le verbe roumain *a trebui* (van Hecke 2007; Rossari et al. 2007), le verbe italien *dovere* (van Hecke 2007; Rocci 2000; Squartini 2004), le verbe espagnol *deber* (Squartini 2004), le verbe danois *måtte / skulle* (Herslund 2003), le verbe estonien *pidama* (Narusk 2014; Treikelder 2014). (Valiukienė 2016 : 9-10)

Ces études, comme nous le mentionnions pour *falloir* en 4.4.5., font appel à des comparaisons entre les langues, car, très souvent, ces autres langues ne possèdent pas de forme impersonnelle déontique comme en français. Les chercheuses tentent alors de voir comment *falloir* est transcrit en contraste avec les formes équivalentes à *devoir*.

Larreya (2004), par exemple, compare les formes déontiques du français à celles de l'anglais :

Il faut ici distinguer d'une part la modalité implicative pure, fondée seulement sur une relation logique (exemple : F *Il faut être fou pour faire ça*, A *You have to be mad to do that*), et d'autre part la modalité implicative associée à une nécessité déontique procédant d'un règlement, d'une loi, etc. (exemple : F *Il faut avoir 18 ans pour avoir le droit de voter*, A *You have to be 18 to be allowed to vote*). Dans le premier cas, les formes employées habituellement sont *IL FAUT* et *HAVE TO*. Dans le second cas, *IL FAUT* et *HAVE TO* sont également très utilisés, mais l'emploi de *DEVOIR* et de *MUST* est assez fréquent : on peut dire (en français) *Vous devez avoir 18 ans pour pouvoir voter* et (en anglais) *You must be at least 18 to be allowed to vote*. (Larreya 2004 : 751)

Pour ce qu'il appelle la modalité implicative – que nous rapprochons dans notre travail de la modalité descriptive – il ne dénote ainsi pas de différence d'interprétation de sens, quel que soit le verbe utilisé. Lorsqu'ils dénotent la valeur modale déontique, les emplois entre *devoir* et *falloir* sont à nuancer. Ainsi, *devoir* « est généralement perçu comme plus solennel ou plus insistant que *IL FAUT* » (Larreya 2004 : 748) et *falloir* est plus usuel « dans des contextes “subjectifs”, soit parce qu'il a l'avantage d'éliminer l'ambiguïté par rapport à une lecture

épistémique [...], soit parce que DEVOIR renverrait à une lecture de “futur de programmation” » (Larreya 2004 : 749). Pourtant, il a été relevé que *falloir* peut aussi revêtir une valeur modale épistémique même si cet emploi est, pour l’instant, moins fréquemment mis en avant dans les recherches que sa valeur modale déontique. La différenciation proposée par Larreya (2004) est donc contestable.

Larreya (2004) n’est pas le seul chercheur à accorder à *devoir* une valeur modale déontique plus forte que pour *falloir*. Ainsi, Dansereau (2016), dans son ouvrage de grammaire destiné aux apprenantes du français d’universités états-uniennes, explique que *falloir* a un sens d’obligation « un peu plus fort » (Dansereau 2016 : 429) que *devoir*. Elle propose les phrases suivantes à titre de comparaison :

62. Je **dois** partir (*must/have to*)

63. Il **faut** que tu partes maintenant (*you have to/must*)

Cependant, étant donné que ces observations ne proviennent pas d’une recherche scientifique mais d’un manuel de langue, elles n’expliquent pas en détail les raisons pour lesquelles l’obligation exprimée par *falloir* est perçue comme plus contraignante que celle exprimée par *devoir*.

De Haan (1997), aussi, estime que *falloir* exprime une valeur modale déontique plus prononcée que *devoir* mais n’en explique pas non plus les raisons :

À bien des égards, le français est similaire à l’italien. Tout comme l’italien, le français dispose essentiellement de deux verbes pour exprimer la modalité forte et faible : *devoir* et *pouvoir*. Les deux sont utilisés dans un sens déontique et épistémique. De plus, le français a la possibilité d’utiliser le verbe *falloir* et la construction impersonnelle *il faut que...* [...] pour exprimer la modalité déontique forte (mais pas la modalité épistémique). (De Haan 1997 : 95)

Avec ce type de remarques, nous pourrions ainsi croire que c’est la forme impersonnelle de *falloir* qui rend sa valeur modale déontique plus forte¹⁰².

La question des degrés et des forces modales ne semble pas souvent faire l’objet de recherches francophones comme c’est le cas pour d’autres langues :

¹⁰² Cette différence pourrait être détaillée avec l’approche de Culioli (1990, 1999a, 1999b) car, comme expliqué en 1.1., elle prend comme point de départ pour l’étude de la modalité l’engagement de la locutrice. Nous ne sommes toutefois pas experte de cette théorie mais nous pouvons imaginer que dans ce cadre, *falloir* serait plutôt classé parmi les modalités de type IV (intersubjectives/déontiques) alors que *devoir* peut aussi être considéré comme faisant partie de la modalité de type II en plus de la IV comme *falloir*.

Dans l'étude de la modalité dans les langues européennes, la nécessité et la possibilité sont parfois considérées en termes de modalité « forte » versus « faible ». Cependant, il est généralement admis qu'il existe plus que deux degrés de distinction sur une échelle de force, puisqu'au moins un degré intermédiaire peut être ajouté [...] Wärensby (2006 : 33) a proposé une « échelle déontique » spécifique allant de « l'obligation » à la « recommandation » puis à la « permission » pour le suédois et l'anglais, reflétant l'autorité décroissante de la locutrice. Narrog (2009a : 90), s'appuyant sur Moriyama (2000), a présenté une échelle à cinq degrés de modalité déontique allant de « l'inévitabilité » à « la permission/exemption », où chaque degré est exprimé par différents marqueurs modaux en japonais. Certains auteurs suggèrent également un gradient ou continuum entre ces degrés (cf. Nuyts 2006 : 5). (Narrog 2016 : 100)

Comme nous pouvons le constater, il existe très peu d'études en français qui comparent ces deux verbes. Souvent, il s'agit de remarques ponctuelles dans des recherches plus vastes sur les marqueurs de modalité, remarques qui peuvent, en plus de cela, se contredire.

4.6.3. Les comparaisons entre *pouvoir*, *devoir* et *vouloir*

En plus des études qui comparent deux verbes modaux, il existe des recherches qui se focalisent sur trois verbes de notre paradigme. C'est le cas notamment des études de Desclés (2003) et Larreya (1997) que nous allons présenter dans cette section. Ces deux auteurs se sont penchés sur *pouvoir*, *devoir* et *vouloir*.

Pour Desclés (2003), les trois verbes possèdent, certes, des effets de sens différents mais ceux-ci seraient tous reliés entre eux :

Les trois valeurs de *pouvoir* sont ancrées sur la valeur fondamentale de *possibilité*. Les trois valeurs de *vouloir* sont ancrées sur les trois valeurs de *pouvoir*, donc sur la valeur de *possibilité*. Les valeurs d'*auto-obligation*, d'*obligation* et d'*ordre* de *devoir* sont ancrées sur la valeur de *volonté* de *vouloir*. (Desclés 2003 : 64)

Les trois valeurs auxquelles Desclés (2003 : 64) fait allusion sont la *possibilité*, la *capacité* et la *permission* pour *pouvoir* ; le *souhait*, la *volonté* et la *demande d'autorisation* pour *vouloir* ; ainsi que la *probabilité*, l'*(auto-)obligation* (l'*ordre*), et la valeur aléthique pour *devoir*. Selon lui, les NS de ces verbes (qu'il appelle *primitives sémantiques*) sont donc la *possibilité* pour *pouvoir*, la *volonté* pour *vouloir* et l'*obligation* pour *devoir*.

Plus que reliées, ces valeurs découlent toutes de *pouvoir*. Ainsi, Desclés (2003) propose le schéma suivant à partir de ses observations :

pourquoi certains effets de sens de ces verbes modaux peuvent se regrouper dans des valeurs modales communes.

Que ce soit les réflexions de Desclés (2003) ou celles de Larreya (1997), elles posent la question de la légitimité d'un NS distinct pour chacun de ces verbes. Si les NS, les valeurs modales et les effets de sens sont tous interdépendants, peut-on véritablement faire une distinction nette entre ce qui est spécifique pour un verbe et pas pour un autre ?

4.6.4. Les comparaisons entre *pouvoir*, *devoir* et *falloir*

Caron et Caron-Pargue (2003) ont comparé *pouvoir*, *devoir* et *falloir*. Elles attestent d'entrée de jeu que ces verbes sont polysémiques mais aussi synonymiques dans certains co(n)textes. Le but de leur étude est ainsi d'étudier comment leurs effets de sens se rejoignent ou se différencient. Pour ce faire, elles font appel aux méthodologies de la linguistique expérimentale. Les autrices ont demandé à des locutrices natives du français de classer différents énoncés dans lesquels les verbes en question apparaissaient selon des paraphrases possibles¹⁰⁴.

Les autrices ont ensuite représenté leurs résultats sous forme d'analyse factorielle des correspondances¹⁰⁵. Il est regrettable de constater que, dans leur article, les autrices ont uniquement représenté de manière visuelle leurs résultats avec *pouvoir* :

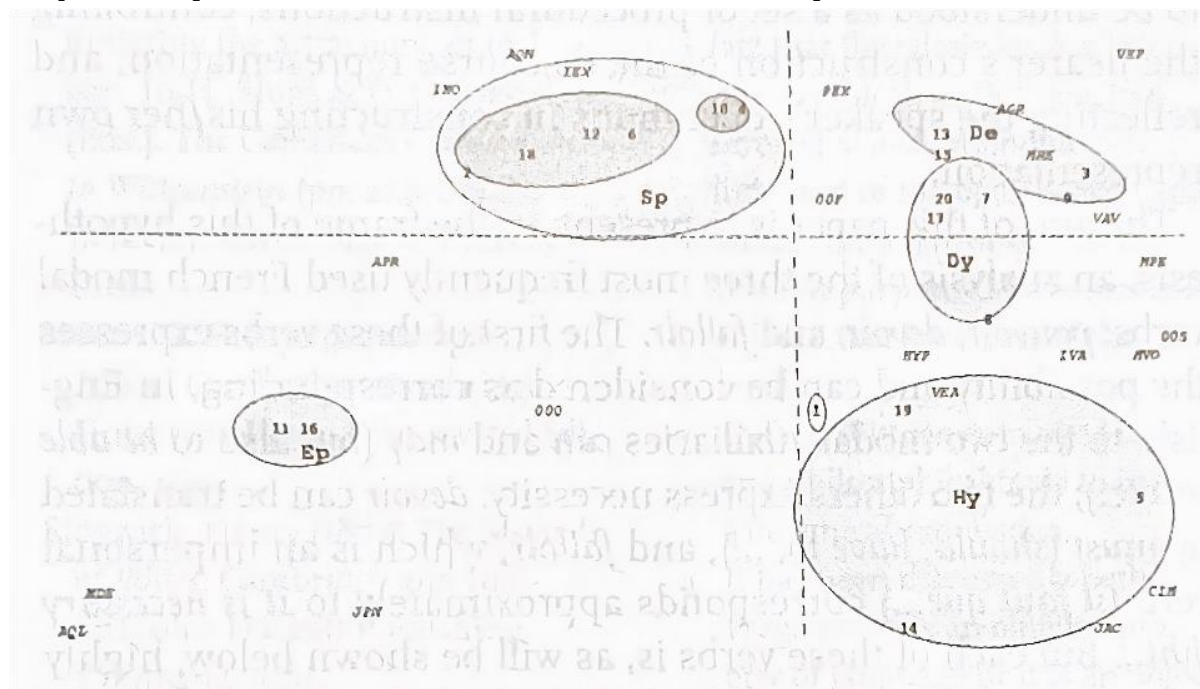


Figure 4 : AFC des paraphrases de pouvoir dans Caron et Caron-Pargue (2003 : 166).

¹⁰⁴ Il est dommage de noter que ces étapes ont été réalisées de manière individuelle pour chaque verbe modal. En effet, on n'a pas laissé aux participantes la chance d'éventuellement créer une catégorie regroupant des effets de sens communs entre *pouvoir* et *devoir* par exemple. L'étude révèle ainsi davantage des caractéristiques individuelles de chacun de ces modaux plutôt que leurs ressemblances et leurs divergences.

¹⁰⁵ Nous reviendrons sur cette méthodologie en 6.4.2.

Dans la figure 4, les chiffres correspondent aux différents énoncés et, les lettres, aux différentes paraphrases possibles. Les formes cylindriques correspondent à des groupements plus généraux que les autres ont ensuite catégorisé avec certaines valeurs modales et effets de sens connus pour ce verbe.

Il est ainsi intéressant de constater que deux phrases se distinguent nettement des autres, soit la 11 et la 16, en formant à elles seules un cluster de la valeur modale épistémique (Ep), dans le cadran en bas à gauche du graphique. Tous les autres effets de sens et valeurs modales se regroupent de manière plus ou moins rassemblée. On y retrouve l'usage sporadique (Sp), déontique (De), hypothétique (Hy) ou dynamique (Dy) de *pouvoir*.

Concernant les analyses qu'elles tirent de cette étude, nous pouvons déjà noter que les effets de sens de *devoir* découlent de ceux de *pouvoir* et qu'il y aurait ainsi une certaine hiérarchie entre ces verbes :

Sur la base des données, nous proposons d'interpréter *devoir* comme exprimant la sélection d'une possibilité, à l'exclusion des autres. En d'autres termes, *devoir* présuppose *pouvoir* : il implique une première étape de construction de l'ensemble des états possibles (impliquée dans la construction des significations de *pouvoir*), la sélection d'un élément de cet ensemble (encore une fois, comme *pouvoir*), puis une opération d'exclusion, écartant toutes les possibilités sauf une. Ainsi, plutôt que d'exprimer une nécessité logique – ce qui pourrait difficilement rendre compte des sens épistémique et “prédicatif” –, il renvoie à une opération de prise de décision, qui peut facilement expliquer les trois sens décrits ci-dessus. (Caron & Caron-Pargue 2003 : 167)

La nécessité exprimée par *devoir* serait ainsi un choix parmi toutes les *possibilités* dénotées par *pouvoir*. Cette réflexion rejoint celles que nous avons relevées dans la section précédente et pose encore une fois la question de la légitimité de distinguer des NS entre les verbes si les caractéristiques sémantiques sont reliées de cette manière.

En plus des corrélations entre les verbes, cette étude a aussi permis de faire émerger des traits spécifiques pour chacun d'entre eux. Si nous rappelons ici les points essentiels de cette étude, nous pouvons dire que

[u]ne première conclusion qui peut être tirée de ces résultats est que le fonctionnement linguistique des verbes modaux n'a que peu à voir avec la logique modale. *Pouvoir* ne correspond pas à l'opérateur modal pur de possibilité, mais développe plusieurs opérations, sur la base de la construction d'un ensemble d'états possibles. *Devoir* n'est pas – comme c'est le cas en logique modale – le terme symétrique par rapport à *pouvoir* : il présuppose la construction d'un ensemble d'états possibles des choses, opère une sélection parmi eux, et exclut tous les éléments sauf un. Et *falloir*, qui serait logiquement équivalent à *devoir*, montre un schéma complètement différent de significations : il établit une relation de dépendance entre l'état ou l'événement décrit et les

conditions dans lesquelles il se produit. Ainsi, chacun de ces trois verbes exprime, non pas une opération logique formelle, mais un ensemble d'instructions pour construire une représentation psychologique. (Caron & Caron-Pargue 2003 : 167)

Ainsi, la dichotomie entre *pouvoir* et *devoir* n'a pas forcément lieu d'être vu que *devoir* semble en partie découler de *pouvoir*. Quant à la ressemblance entre *devoir* et *falloir*, celle-ci serait également à réexaminer vu que *falloir* peut exprimer des effets de sens et des valeurs modales bien différentes de celles de *devoir* selon cette étude.

Synthèse sur la sémantique des verbes modaux

Tout au long de ce chapitre 4, nous avons montré que les verbes modaux sont dotés de caractéristiques sémantiques et ne peuvent ainsi pas être classés comme de simples auxiliaires ou semi-auxiliaires au même niveau qu'*être* et *avoir*. Certes, il existe des usages dans lesquels, par exemple, *devoir* ou *vouloir* semblent véhiculer uniquement un usage périphrastique, au même plan qu'*aller*, mais nous avons pu établir qu'ils étaient loin de se réduire à ce type d'emploi.

Cet état de l'art que nous avons proposé ici n'est évidemment que partiel mais il a le mérite de déjà faire ressortir de nombreuses propriétés concernant la sémantique de *pouvoir*, *devoir*, *falloir* et *vouloir*.

Tout d'abord, nous avons réussi à mettre en exergue des notions importantes qui méritent d'être différenciées lorsque nous pratiquons des recherches en sémantique des marqueurs modaux. Nous avons identifié les concepts suivants qui permettent de comparer les verbes modaux sur le plan sémantique :

- I. Le noyau sémantique : c'est la notion centrale à partir de laquelle découlent tous les effets de sens et valeurs modales cotextuelles que les verbes modaux peuvent véhiculer.
- II. Les effets de sens : ce sont les notions véhiculées par les verbes selon le co(n)texte.
- III. Les valeurs modales : ce sont les modalités exprimées par les verbes modaux. Elles sont généralement de l'ordre de la modalité épistémique, déontique, descriptive ou encore boulique pour les quatre verbes que nous étudions ici.
- IV. Les phénomènes de grammaticalisation : ce sont les occurrences dans lesquelles le noyau sémantique n'est pas ou peu perceptible.

C'est pour les points II et III que nous pouvons considérer que les verbes modaux sont polysémiques. En effet, tous peuvent transcrire différents effets de sens et valeurs modales.

En ce qui concerne les NS des verbes modaux, un consensus semble se dégager parmi les chercheuses en ce qui concerne les verbes *pouvoir* et *devoir*. Il est largement admis que *pouvoir* exprime la *possibilité* tandis que *devoir* renvoie à la *nécessité*. Ces NS sont plus couramment acceptés que d'autres car ils englobent un large éventail d'effets de sens et de valeurs modales associés à ces verbes. Toutefois, les NS de *falloir* et *vouloir* sont moins étudiés à ce sujet. Malgré cela, en nous appuyant sur les travaux existants, nous avons observé que pour le verbe *falloir*, il est possible d'attribuer un NS de *contrainte* ou de *nécessité*, tandis que pour *vouloir*, un NS d'*intention* ou de *volonté* peut être envisagé.

Notons que parfois, ces notions ne sont pas forcément distinctes. Ainsi, pour *pouvoir*, les autrices admettent un NS de *possibilité* mais c'est également un effet de sens pour d'autres. Dès lors, nous pouvons questionner la légitimité de différencier ces niveaux sémantiques voire même la pertinence d'attribuer un NS pour les verbes modaux si celui-ci est aussi un effet de sens ou une valeur modale co(n)textuelle.

Si nous admettons un NS pour ces verbes, il arrive que celui-ci ne soit pas facilement récupérable dans certains usages. Nous pouvons, dans ces cas-là, parler de phénomènes de *grammaticalisation*.

Nous avons aussi relevé le fait que les effets de sens et les valeurs modales, voire les emplois grammaticalisés, dépendent du cotexte, plus particulièrement du type de sujet qui régit ces verbes et de la complémentation verbale sous forme d'un verbe à l'infinitif. Ces variations n'ont été que très peu étudiées à l'aide d'outils statistiques et les observations que nous avons regroupées ici sont surtout le résultat d'exemples manipulés par les chercheuses.

En plus des réflexions individuelles sur chacun de ces verbes, nous avons aussi pris le temps de relater les différentes études qui comparent ces verbes. Il est intéressant de noter qu'elles ne sont pas nombreuses et qu'elles ne prennent pas en compte les quatre verbes étudiés ici alors que, comme nous l'avons relevé, ils possèdent des effets de sens proches et des valeurs modales communes. Certes, chacun d'entre eux possède des particularités, ce qui le différencie des autres, mais les points communs qu'ils partagent n'ont jamais été traités. Il est ainsi intéressant de comparer plusieurs verbes pour faire émerger les propriétés sémantiques propres à chacun d'eux. Les étudier seuls ne permet pas toujours de faire ressortir leurs particularités sémantiques.

Ces réflexions théoriques nous permettent d'aboutir aux différentes lacunes que cette thèse propose de combler.

5. Objectifs de la présente thèse

Cette thèse s'attache à examiner les caractéristiques sémantiques des verbes modaux en s'appuyant sur les analyses des approches qualitatives largement prévalentes dans les études existantes à leur sujet. Nous avons, par exemple, pu lister qu'il existe pour chacun de ces verbes, à savoir *pouvoir*, *devoir*, *falloir* et *vouloir*, des NS, des effets de sens ainsi que des valeurs modales plus ou moins acceptées mais aussi des co(n)textes de désémantisation et/ou de grammaticalisation.

Nous avons pu établir pour chacun des verbes modaux que le co(n)texte influence les effets de sens et les valeurs modales de ces verbes, voire même les phénomènes de grammaticalisation dont ils font l'objet. C'est plus particulièrement le sujet grammatical ainsi que la complémentation verbale sous forme de verbe à l'infinitif qui semblent être primordiaux pour déterminer l'effet de sens et/ou la valeur modale du verbe. Cet intérêt n'est pas étonnant lorsqu'on sait qu'en 1957 déjà Firth écrivait : « You shall know a word by the company it keeps » (Firth 1957 : 11).

Malheureusement, cette approche n'a été menée que de manière intuitive pour les verbes modaux : les chercheuses ne se sont pas basées sur des outils quantitatifs pour étudier l'influence du cotexte pour les interprétations sémantiques de *pouvoir*, *devoir*, *falloir* et *vouloir*. A notre connaissance, cette influence du cotexte pour les verbes modaux en français n'a été mise en lumière qu'à partir de manipulations d'exemples, ou, dans le meilleur des cas, d'observations d'exemples authentiques tirés de textes divers et variés. Pourtant, il existe des outils qui permettent de calculer de manière statistique les associations privilégiées entre plusieurs unités lexicales.

La recherche de cette thèse repose sur des outils de la linguistique de corpus dont les fonctionnements sont présentés dans le chapitre suivant. Le but de cette étude est de comparer *pouvoir*, *devoir*, *falloir* et *vouloir* en observant leurs associations privilégiées avec les sujets grammaticaux et leur complémentation verbale sous forme de verbe à l'infinitif. Cette analyse permettra de faire apparaître les similarités et différences entre ces quatre verbes modaux et donc leurs singularités les uns par rapport aux autres¹⁰⁶.

¹⁰⁶ Nous pensons, à l'instar de l'analyse sémique de Pottier (1964) et des réflexions sur le langage de Wittgenstein (1961), que c'est par comparaisons entre unités lexicales aux NS et effets de sens proches, que nous pouvons faire ressortir les spécificités de chacune. Les recherches comparatives que nous avons citées précédemment confirment également l'importance des mises en parallèle de plusieurs items afin de faire ressortir des propriétés, certes communes, mais aussi particulières à chacune des unités lexicales étudiées.

Notre recherche sera la première à comparer ces quatre verbes qui, comme nous l'avons relevé, possèdent des propriétés sémantiques communes. Il semble ainsi pertinent de les confronter sur ce plan.

Nos questions de recherches sont donc les suivantes :

- i. Quelles sont les similarités et les différences sémantiques de *pouvoir*, *devoir*, *falloir* et *vouloir* que l'on peut observer à l'aide de leur environnement privilégié ?
- ii. Quelles sont les particularités sémantiques pour chacun de ces verbes qui ressortent des associations statistiquement significatives observables à l'aide des outils de la linguistique de corpus ?
- iii. Est-ce que ces propriétés observées sont stables selon le texte dans lequel apparaît le verbe modal ?
- iv. Ces caractéristiques varient-elles selon le temps et/ou le mode auquel le verbe est conjugué ?
- v. En somme, qu'est-ce que les méthodes quantitatives permettent de confirmer ou d'infirmer des informations déjà disponibles sur les verbes modaux dans les recherches qualitatives ? Ou apportent-elles des informations nouvelles sur ces verbes encore méconnues aujourd'hui ?

Nos hypothèses et propositions méthodologiques pour les tester, fondées sur ce que nous savons désormais des verbes modaux, sont les suivantes :

- i. Certaines propriétés des verbes modaux peuvent être vérifiées à l'aide d'outils quantitatifs comme par exemple les associations des co(n)textes grammaticalisés comme *vouloir dire*, *faut croire/voir/savoir*, etc.
- ii. Nous pensons que certaines similarités et différences entre les verbes modaux dénombrées dans la littérature pourront être confirmées. Ainsi, nous estimons, par exemple, que *devoir* et *falloir*, les deux verbes dénotant principalement les valeurs modales déontiques, peuvent avoir des cotextes d'utilisation proches.
- iii. *Pouvoir* et *devoir*, les deux verbes prototypiques des modalités épistémiques et déontiques, seront très différents dans leurs comportements. En effet, même s'ils sont souvent comparés sur le plan épistémique, leurs NS respectifs sont trop distincts pour qu'ils soient utilisés dans des co(n)textes similaires.
- iv. *Vouloir*, qui n'est pas toujours considéré comme un verbe modal selon les autrices, se distingue des autres verbes par son co(n)texte. Le fait qu'il soit parfois considéré comme un verbe sémantiquement plus riche que les trois autres fait qu'il aura des

comportements qui le différencient de ceux-ci. Ses nombreux usages grammaticalisés peuvent également faire ressortir des particularités cotextuelles.

- v. Les associations privilégiées de ces verbes modaux nous permettront de confirmer ou alors de proposer d'autres NS pour chacun de ces verbes.

Ces hypothèses seront testées à l'aide des outils de la linguistique de corpus que nous présentons au chapitre suivant.

6. Méthodologie

Nous avons pu soulever le fait que *pouvoir*, *devoir*, *falloir* et *vouloir* possèdent des caractéristiques communes mais aussi divergentes. Celles-ci peuvent ressortir notamment à l'aide du cotexte dans lequel ces verbes se situent. Dans ce chapitre, nous allons présenter comment étudier ce cotexte pour faire émerger les particularités sémantiques des verbes modaux. Pour ce faire, nous allons exposer les divers outils méthodologiques que nous empruntons tant à la tradition anglophone de linguistique de corpus qu'à la tradition francophone de la textométrie.

Nous commencerons par retracer l'histoire de la linguistique de corpus et de la textométrie, avant de nous attarder sur leurs différences et leur complémentarité dans le cadre de notre recherche (6.1). Nous définirons ensuite les différentes notions tirées de ces méthodologies mobilisées tout au long de la présentation de nos résultats, telles que la différence entre les approches corpus *based* ou *driven* (6.2.), la notion de corpus (6.3.1.), de fréquence (6.3.2.), la cooccurrence (6.3.3.) et l'empan (6.3.4.). Il sera également question de la représentation de ces résultats générés avec l'indice de spécificité (6.4.1.) – qui permet d'étudier le profil combinatoire d'un item en particulier – mais aussi les méthodes d'analyses factorielles des correspondances (6.4.2.) ainsi que sur l'analyse arborée (6.4.3.) – deux méthodes qui permettent de comparer les items entre eux avec des représentations graphiques. Pour finir, nous présenterons les corpus (6.5.1.) et les divers logiciels dont nous faisons usage, c'est-à-dire TXM (6.5.2.1.) et Hyperbase (6.5.2.2.).

Le but de cette partie méthodologique sera ainsi de démontrer l'utilité des approches quantitatives pour l'étude des propriétés sémantiques des verbes modaux afin de faire émerger leurs ressemblances et leurs divergences.

6.1. La linguistique de corpus et la textométrie

Les outils que nous utilisons dans cette recherche viennent de deux approches quantitatives quelque peu différentes mais complémentaires : la linguistique de corpus et la textométrie.

Les outils de la linguistique de corpus sont apparus dans les années 1950 avec l'arrivée des ordinateurs et des textes informatisés (McEnery & Hardie 2013 : 728) mais ce n'est qu'en 1984 qu'apparaît pour la première fois cette notion comme titre d'un ouvrage collectif qui regroupait « un ensemble de travaux britanniques, scandinaves et néerlandais sur des corpus informatisés

de l'anglais » (Léon 2008 : 12). Ces méthodes se focalisent sur l'usage (Léon 2008 : 13) et on parle ainsi

de linguistique de corpus (Habert et al. 1997 ; Biber et al. 1998 ; McEnery & Wilson 2001 ; Cheng 2011 ; McEnery & Hardie 2012) à partir du moment où ce sont des données langagières attestées - et pas seulement la réflexion et l'intuition du linguiste - qui permettent d'étayer des théories et des hypothèses linguistiques. (Poudat & Landragin 2017 : 14)

Cette vision s'oppose alors à celle de Chomsky dont les méthodes de recherche reposaient sur l'intuition de la locutrice native et sa capacité à produire un nombre infini d'énoncés (Chomsky 2002). Ainsi, se reposer sur des données « finies », ne nous dit « peu ou rien sur la compétence, précisément parce que le nombre de phrases possibles dans une langue est infini, et toute collection de données de performance ne peut représenter qu'un sous-ensemble fini (et biaisé) de la langue » (McEnery & Hardie 2013 : 730). Nait alors un clivage entre linguistes chomskyennes et celles qui travaillent sur des données attestées. L'intérêt premier pour les linguistes de corpus est « d'accéder à la description de phénomènes textuels qui présentent un grand intérêt une fois mis en évidence et dont il aurait été difficile de cerner les contours *a priori* » (Habert et al. 1997 : 184) ce qui n'est pas possible avec l'intuition seule.

Travailler sur des données authentiques en grand nombre a été possible pour les linguistes grâce au développement informatique des dernières décennies. En effet, « [l']accès actuel à de vastes ensembles de textes sous forme électronique a été une condition décisive pour le développement [de la linguistique de corpus] » (Pincemin 1999 : 26). Ce statut particulier de la linguistique de corpus qui met en avant les outils lui confère un statut particulier par rapport à d'autres branches de la linguistique :

Contrairement à des domaines comme la phonologie, l'étude de la variation sociale dans la langue ou l'analyse critique du discours, la linguistique de corpus ne concerne pas directement l'étude d'un aspect particulier de la langue. Son objectif porte plutôt sur un ensemble de procédures, ou méthodes, pour étudier la langue. La linguistique de corpus est donc une méthodologie qui peut être appliquée à différents domaines d'étude au sein de la linguistique. (McEnery & Hardie 2013 : 727)

Ainsi, lorsqu'une chercheuse se définit comme linguiste de corpus cela ne nous renseigne pas sur les études qu'elle mène mais avant tout sur les outils qu'elle utilise. Cependant, une autre vision de la linguistique de corpus est de la considérer, non pas seulement comme une méthodologie, mais comme objet d'étude en tant que tel :

Il convient toutefois de noter que cette vision méthodologique de la linguistique de corpus n'est pas la seule approche : une vision alternative du domaine, développée par l'école néo-firthienne sous la direction de John Sinclair, conceptualise la linguistique de corpus non pas comme une méthodologie,

mais comme ayant un statut théorique en soi – reliant les résultats de l’analyse de corpus à une compréhension de la langue où les mots, leurs significations dans le contexte discursif et leurs patterns de collocation sont centraux. (McEnery & Hardie 2013 : 728)

Dans notre thèse, nous considérons la linguistique de corpus non seulement comme un outil d’analyse, mais également comme un moyen de consolider les connaissances méthodologiques dans ce domaine. Son utilisation contribuera à éclairer les chercheuses sur cette approche et facilitera l’adoption de techniques similaires par d’autres, tout en offrant la possibilité d’affiner ces méthodes pour une utilisation plus efficace et précise.

La linguistique de corpus a permis aux sciences du langage une première ouverture sur les méthodes quantitatives. Depuis, d’autres méthodologies ont été développées. Nous souhaitons mettre ici en évidence que nous utilisons tant des outils de la linguistique de corpus que de la textométrie, méthodologie employée majoritairement dans les régions francophones, et plus précisément en France. La textométrie « se caractérise par la place centrale du texte tout au long de l’analyse » (Mayaffre et al. 2019 : 102) et les chercheuses sont ainsi davantage issues du domaine de l’analyse du discours. La linguistique de corpus, elle, « [...] s’attache plutôt à l’analyse d’un marqueur linguistique décrit en usage [...] » (Mayaffre et al. 2019 : 102). Dans la linguistique de corpus,

[I]la question du texte et de ses catégorisations reste ainsi secondaire alors même qu’elle est centrale pour la textométrie, ce qui entraîne des différences importantes dans le développement des méthodes et des outils. Par exemple, l’analyse des correspondances, centrale pour la textométrie, est absente des outils classiques de la linguistique de corpus. (Mayaffre et al. 2019 : 102)

Une autre différence fondamentale entre les deux disciplines concerne également la taille des corpus. En effet, « la linguistique empirique britannique, vise à décrire les usages de la langue à travers l’étude de “très grands corpus” » (Sitri & Barats 2017b : 11) de plusieurs millions de mots alors que les études textométriques peuvent se baser sur des corpus plus restreints. Dans sa thèse, Montrichard (2020) travaille sur un corpus de presse de tranchées représentant 541’825 occurrences dont la taille

est très restreinte par rapport aux vertigineux corpus avoisinant les dizaines de millions de mots que d’aucuns traitent en textométrie (Patin et al., 2016). Sa taille est toutefois supérieure à un certain nombre de corpus donnant lieu à des analyses qui n’en sont pas moins pertinentes et solides [...] Il n’y a pas de normes quantitatives clairement définies en la matière : ce sont les textes, les hypothèses de travail et les objectifs de l’étude qui induisent la taille du corpus. Par exemple, les travaux visant à l’analyse des textes du chanteur-parolier Hubert-Félix Thiéfaine (Beucher-Marsal & Kerneis, 2016) ont porté sur un corpus d’un peu plus de 40 000 occurrences. (Montrichard 2020 : 81)

La textométrie a comme ancêtre la *statistique lexicale*, dont l'objet d'étude était les textes littéraires et le vocabulaire usité dans ceux-ci (Sitri & Barats 2017b : 10) tout comme la *lexicométrie*, proche de la statistique lexicale mais qui

s'en différencie par le fait qu'elle s'intéresse non pas aux particularités du style d'un auteur mais aux régularités d'un discours, en les mettant en relation avec des déterminations idéologiques ou des positionnements sociaux : c'est ainsi que le Laboratoire de lexicologie politique de l'ENS de Saint-Cloud (ci-après Laboratoire de Saint-Cloud) étudie les tracts de Mai 68 ou les résolutions des congrès syndicaux. La lexicométrie est ainsi très proche de l'AD¹⁰⁷, mais les deux disciplines se distinguent par le choix des observables : à ses débuts la lexicométrie s'intéresse avant tout au lexique tandis que l'AD, sans pour autant négliger l'unité lexicale, faisait de la syntaxe (les relatives, les nominalisations) son objet d'observation privilégié. À l'heure actuelle, les logiciels de lexicométrie évoluent de plus en plus vers la prise en compte du texte et l'on parle alors plus volontiers de textométrie. (Sitri & Barats 2017b : 10-11)

La textométrie est aussi arrivée après la linguistique de corpus. Nous pourrions résumer cette évolution en disant que ce sont d'abord les mots qui ont été étudiés de manière quantitative puis les textes (Poudat & Landragin 2017 : 17).

Il est toutefois important de rappeler que « [m]algré cette différence d'objets et de programme de recherche entre linguistique de corpus et analyse de discours, existent un certain nombre de convergences » (Jacques 2016 : 90) entre ces deux courants car « nombre d'articles de recherche en analyse de discours sont le fait de linguistes et qu'ils y mobilisent des concepts linguistiques » (Jacques 2016 : 90). Les outils et les notions peuvent être similaires, par exemple « l'attention prêtée au contexte dans l'interprétation du sens » (Sitri & Barats 2017b : 11) est importante tant pour les chercheuses revendiquant plutôt appartenir à la linguistique de corpus qu'à la textométrie. Les différences entre ces deux approches peuvent être plus ou moins marquées selon la perspective des chercheuses. Dans nos recherches, nous montrons, par exemple, comment les logiciels conçus à l'origine pour la textométrie peuvent être adaptés à l'analyse de marqueurs linguistiques tels que les verbes modaux. De plus, en étudiant ces marqueurs dans deux types de textes différents, nous mettons en évidence les variations d'usage possibles selon le genre textuel.

En résumé, les personnes adoptant une approche textométrique se concentrent sur les caractéristiques d'un texte, d'un genre ou d'une autrice, tandis que les linguistes de corpus étudient les spécificités d'un mot, d'une suite de mots ou d'une catégorie particulière d'unités lexicales. Dans notre thèse, ces deux approches se rejoignent puisque nous analysons les

¹⁰⁷ Pour *analyse du discours*.

propriétés de quatre verbes modaux dans deux types de textes distincts. Notre objectif n'est pas de proposer un éclairage nouveau sur les genres textuels, mais de fournir une analyse des verbes modaux à travers le prisme des genres textuels.

6.2. Une approche corpus-*based* ou/et corpus *driven*

Une des différenciations majeures sur la façon d'entrevoir la recherche avec des corpus est d'utiliser soit une méthode *based* ou *driven*. La méthode *based*, comme son nom l'indique, « se base » sur des hypothèses préconstruites et les vérifie dans un corpus. Pour le reformuler, cette première méthode

consiste à utiliser les données présentes dans un corpus comme socle pour une théorisation linguistique. [...] Il s'agit de la démarche déductive. Elle est focalisée, en ce sens qu'elle implique plus ou moins le fait que le chercheur a déjà en tête une hypothèse linguistique avant d'aller regarder ce qu'il en est en corpus. Elle sous-entend également que le résultat obtenu suffira à confirmer ou à infirmer l'hypothèse, et qu'en tout cas l'exploitation et l'exploration du corpus ne vont pas plus loin : l'extraction d'exemples authentiques se suffit à elle-même. (Poudat & Landragin 2017 : 31)

Avec cette approche, la chercheuse confirme ou infirme ses hypothèses, et tente de démontrer que son intuition se confirme dans des données attestées. Cette démarche vise ainsi à valider « les formes et structures linguistiques dérivées de la théorie linguistique » (Biber 2015 : 196). Cette première méthode fait partie de notre recherche puisque nous observons si les réflexions présentées dans les études qualitatives sur les verbes modaux concernant leurs cotextes privilégiés se confirment dans des données authentiques et grâce à des calculs statistiques. La pratique corpus *driven*, quant à elle, consiste à se laisser totalement guider par les données du corpus :

La théorie se construit en quelque sorte sous la direction des données. Il s'agit de la démarche inductive. Elle implique plus ou moins le fait que le chercheur ne cherche pas à formuler d'hypothèse avant de regarder les données. On attend d'elle qu'elle permette de découvrir des régularités, des patrons linguistiques, des observations qui feront progresser le chercheur dans son champ linguistique. (Poudat & Landragin 2017 : 30)

Ainsi, les linguistes utilisant cette méthode n'ont *a priori* pas d'idée précise concernant leur objet de recherche. Cependant, les recherches exclusivement corpus *driven* sont rares car les linguistes n'ont rarement aucune hypothèse sur les données. Cela impliquerait de soutenir que

seule l'existence des mots est donnée *a priori*, tandis que des concepts tels que la 'phrase' et la 'proposition' n'ont pas de statut *a priori*. Au contraire, ce sont les motifs de cooccurrence entre les

mots, découverts à partir de l'analyse de corpus, qui servent de fondement aux descriptions linguistiques ultérieures. (Biber 2015 : 196)

Les chercheuses adoptant cette approche sont particulièrement attentives à la fréquence (Biber 2015 : 210) puisqu'il s'agit simplement de lister des phénomènes récurrents qui ne sont pas forcément attendus.

Les deux approches peuvent également être différenciées par le fait que la première utiliserait les corpus comme « simple réservoir d'exemples destiné à tester ou vérifier des positions théoriques existantes, [alors que] l'option "corpus-driven" [...] postule qu'aucune position théorique a priori ne préside aux observations sur corpus, la théorie étant induite du corpus » (Léon 2008 : 12)¹⁰⁸.

Dans une certaine mesure, il est également possible de lier ces deux visions dans une seule et même recherche. En effet, « [l]es deux types d'approches inductive et déductive ne sont bien sûr pas exclusifs et il est même parfaitement d'usage de combiner les deux dans des parcours méthodologiques bien pensés » (Poudat & Landragin 2017 : 31). Ainsi, nous adoptons dans cette thèse une démarche mixte : nous avons formulé au préalable des hypothèses à la suite des études préexistantes sur les verbes modaux. Nous allons observer les sujets mais aussi les complémentations verbales sous forme de verbe à l'infinitif car ces deux éléments semblent influencer grandement sur les effets de sens et valeurs modales de ces verbes. En ceci, nous nous insérons dans une démarche corpus *based*. Toutefois, notre démarche est également *driven* car nous approfondissons les pistes qui ressortiront à travers les premiers résultats.

6.3. Les diverses notions mobilisées dans cette recherche

Dans cette section, nous allons présenter les différentes notions mobilisées tout au long de l'analyse. Certains concepts semblent très intuitifs même pour les chercheuses non initiées à ce type de méthode. Cependant, il nous semble important de les définir ici et d'explicitier la manière dont nous allons les déployer dans notre recherche afin de rendre la lecture de nos résultats la plus limpide possible.

¹⁰⁸ Cette démarche semble plus applicable à une étude textométrique pure. Ainsi, une chercheuse peut commencer à regarder les mots les plus fréquents de son corpus, par exemple selon l'autrice, les dates ou une autre composante du corpus avant de creuser une piste en particulier.

6.3.1. Les corpus

La notion de corpus est généralement décrite selon les principes de Sinclair, rappelés par Habert et al. (1997) : « [...] le mot *corpus* dans une acception restreinte empruntée à J. Sinclair (1996, p.4) : “[...] est une collection de données langagières qui sont sélectionnées et organisées selon des critères linguistiques explicites pour servir d’échantillon au langage.” » (Habert et al. 1997 : 11). Depuis, cette définition a été affinée. Yoon et Gries (2016) intègrent, par exemple, des notions liées à l’informatique et au traitement des données en tant que telles :

La notion de corpus est une catégorie prototype et le prototype d’un corpus est une collection de fichiers contenant du texte et/ou des discours transcrits censés être représentatifs et équilibrés pour une langue, une variété, un registre ou un dialecte donné : souvent, les fichiers contiennent non seulement la langue écrite ou parlée (souvent encodée en UTF-8 pour couvrir différentes orthographe), mais aussi des annotations supplémentaires (souvent des informations de partie du discours, mais aussi morphologiques, sémantiques, ou d’autres informations sous forme d’annotations XML). Les corpus diffèrent principalement par leur taille (de quelques récits narrés dans une langue sous-documentée voire déjà éteinte à des corpus contenant plusieurs milliards de mots) et par le degré de naturalité des données qu’ils contiennent (allant du dialogue complètement spontané entre deux locuteurs à des situations expérimentales hautement contraintes). (Yoon & Gries 2016 : 2)

Les autrices insistent, comme Habert et al. (1997), sur la fonction « prototype du langage », mais ajoutent aussi les dimensions informatiques aux données, comme l’encodage et les annotations, éléments qui font qu’un corpus puisse être exploitable dans tel ou tel logiciel. Elles soulignent également l’importance de la taille des corpus¹⁰⁹ et des types de textes qui les composent.

La chose la plus importante à retenir est le fait que le corpus « [...] est une collection de **productions langagières attestées** [...] » (Poudat & Landragin 2017 : 11). Il ne se compose pas d’exemples mais

contient des données issues de *situations communicatives naturelles*, ce qui signifie que lors de la production des données linguistiques dans le corpus, celles-ci n’ont pas été produites uniquement dans le but d’être intégrées à un corpus, et/ou que la production des données linguistiques a été aussi peu influencée que possible par la collecte de ces données. (Gries & Berez 2017 : 380)

¹⁰⁹ Rappelons que le critère de la taille a surtout son importance dans une approche anglo-saxonne de la linguistique de corpus.

C'est donc sur des données authentiques de la langue¹¹⁰ que reposent les résultats obtenus à l'aide des outils de la linguistique de corpus afin de faire émerger ou de confirmer des phénomènes linguistiques.

Un corpus est par nature constitué de données attestées et permet d'étudier la langue telle qu'elle est, dans un contexte et moment précis. Les corpus sont également exploitables de manière informatique et les textes qui le composent respectent donc le format du logiciel ou de la plateforme dans lesquels ils sont insérés.

En plus de ces critères « formels » qui concernent en grande partie l'authenticité des données, il convient d'ajouter également des normes davantage liées à la recherche en tant que telle. En effet,

le corpus n'est pas un simple recueil de textes qui seraient à disposition et qu'il suffirait de compiler.

Il est construit en fonction des questions et des hypothèses de recherche, voire même en fonction de la conception que l'on a de l'analyse du discours, ou encore des outils ou des catégories que l'on utilise. (Sitri & Barats 2017a : 41)

Pincemin (1999 : 27) a évoqué trois types de conditions dont les chercheuses doivent tenir compte lorsqu'elles construisent un corpus :

- I. Conditions de signifiante : « [u]n corpus est constitué en vue d'une étude déterminée [...] portant sur un objet particulier » (Pincemin 1999 : 27) et doit également présenter un angle de vue précis et non pas les multiplier. En d'autres termes, il serait inadéquat de composer un corpus alliant des articles de presse traitant des élections présidentielles à des extraits de brochures touristiques. Effectivement, tant le sujet que la forme ne présenteraient aucune concordance.
- II. Conditions d'acceptabilité : le corpus doit représenter fidèlement la langue et les données externes qui le composent¹¹¹. L'ensemble de ces éléments doit être représenté de manière précise et régulière, tout en restant en adéquation avec l'analyse produite. Ainsi, dans un

¹¹⁰ Cette authenticité peut parfois être questionnée. En effet, un corpus comme le *Corpus Français* de l'Université de Leipzig (Biemann et al. 2007), consultable en ligne (https://corpora.uni-leipzig.de/fr?corpusId=fra_mixed 2012), est construit sur un principe de collections de phrases issues de textes authentiques et non pas les textes dans leur intégralité. L'authenticité peut-elle encore être perçue lorsque nous traitons la langue dans des phrases isolées ? Cette authenticité peut aussi être interrogée avec un corpus comme OFROM composé en partie « d'entretiens guidés à dominante monologique, dans lesquels l'interviewé [...] était sollicité pour répondre à des questions nécessitant des réponses plus ou moins longues posées par l'intervieweur » (Avanzi et al. 2016 : 3). Ces entretiens ont été menés expressément pour construire une partie de ce corpus. Pourtant, il semble bien que selon les définitions que nous avons soulevées ici qu'un corpus de textes devrait se fonder sur des textes qui n'ont pas été créés spécifiquement pour une exploitation linguistique. Il est ainsi nécessaire de préciser à chaque fois d'où viennent les données des corpus mobilisés.

¹¹¹ Ces données externes font référence aux métadonnées comme le nom des autrices des articles, la date de parution, etc.

corpus de presse, il serait inadéquat que seuls certains articles qui composent ce corpus soient annotés avec la date de parution et d'autres non.

- III. Conditions d'exploitabilité : les textes qui composent le corpus doivent être homogènes et en nombre suffisant afin de « pouvoir repérer des comportements significatifs (au sens statistique du terme) » (Pincemin 1999 : 27). Il ne convient ainsi pas de mélanger des types de discours distincts, tout comme il est nécessaire d'avoir un corpus avec un nombre d'occurrence significatif afin de mener à bien les analyses prévues.

Afin que la recherche et les résultats qui découlent des corpus soient adéquats, il est nécessaire d'avoir une bonne « connaissance de la constitution du corpus (dont la perspective qui a orienté la sélection des textes) et de ses limites, de ses “défauts” » (Pincemin 2012 : 16). Nous présenterons dans la section 6.5.1. nos corpus en les mettant en lien avec ces conditions tout en signalant leurs imperfections et faiblesses.

6.3.2. La fréquence

La fréquence est une notion assez intuitive. Il s'agit simplement du nombre de fois qu'un terme, ou une série de mots, apparaît dans le corpus. Cependant, nous pouvons distinguer la manière dont nous l'employons dans les recherches faisant appel aux outils de la linguistique de corpus. Ainsi, il existe les fréquences absolues, ce que Leblanc et al. (2017) appellent « occurrences » (Leblanc et al. 2017 : 138). Lorsqu'on parle de fréquence absolue, cela : « [...] inclurait comme cas limite des études basées sur l'observation de la présence (et de la manière) de quelque chose, mais peut également inclure des cas où les fréquences ou les probabilités sont utilisées pour classer les mots dans des constructions, etc. » (Yoon & Gries 2016 : 2). Cette mesure, en tant que telle, n'est pas souvent utilisée car elle permet simplement d'attester d'un fait linguistique ou de sa non-existence. Toutefois, c'est une première donnée intéressante à prendre en compte, ne serait-ce que pour déterminer si la thématique de notre étude est pertinente ou non par sa simple fréquence¹¹².

La notion que l'on emploie habituellement lorsqu'on parle de fréquence en linguistique de corpus est davantage la *fréquence relative*. Pincemin (2012) explique qu'il faut « rapporter cette fréquence à la longueur de la partie, de sorte à obtenir une *fréquence relative* » (Pincemin 2012 : 20). Dans notre cas, nous étudions la fréquence relative d'un mot ou d'une suite de mots par

¹¹² Dans le cas d'une approche corpus *driven* il semble, par exemple, nécessaire de passer par cette simple étape des fréquences absolues afin d'apprendre à connaître le corpus utilisé.

rapport à l'ensemble des autres mots du corpus, et non par rapport à des sections spécifiques d'un texte. Cette méthode permet également d'obtenir une fréquence relative.

6.3.3. La cooccurrence

« You shall know a word by the company it keeps » (Firth 1957 : 11) est le principe même sur lequel se fondent les premières recherches sur corpus. Les chercheuses se sont penchées sur l'entourage d'un mot pour pouvoir le connaître, le définir et en décrire les propriétés. Il existe plusieurs terminologies¹¹³ pour nommer cet entourage :

Ce phénomène se décline avec le vocabulaire suivant : **cooccurrences**, c'est-à-dire la présence simultanée de mots ou groupes de mots dans un même extrait de taille restreinte, ce qui indique une proximité non seulement spatiale mais aussi sémantique ; **segments répétés**, c'est-à-dire l'apparition de plusieurs occurrences d'un même groupe de mots consécutifs (Lafon & Salem 1983) ; ou encore **motifs textuels**, notion qui étend et généralise les deux précédentes en considérant des successions récurrentes de mêmes unités textuelles [...]. (Poudat & Landragin 2017 : 26)

La notion qui correspond le plus à notre façon d'envisager cette étude est celle de *cooccurrence*. Sa définition la plus rudimentaire est « [...] l'association statistiquement significative de deux items (en général deux mots) dans une fenêtre déterminée du texte [...] » (Mayaffre 2014 : 4) ou encore « la coprésence régulière de deux unités linguistiques dans une fenêtre contextuelle choisie » (Poudat & Landragin 2017 : 201) . Nous l'opposons aux *collocations*,

[...] définies comme des associations lexicales privilégiées et sémantiquement compositionnelles (ex : *tristesse infinie*, *pertes abyssales*, *jouer un rôle*, etc.) [qui] constituent désormais une notion essentielle dans les approches contemporaines de la phraséologie. (Tutin 2013 : 47).

Pour résumer, le terme *cooccurrence* est plus général car il « [...] désigne [...] le plus souvent des associations statistiques (critère de fréquence) dont on constate les relations textuelles (plus que phrastiques), sans distinction de nature phraséologique, sémantique, thématique, rhétorique, stylistique, etc. » (Mayaffre 2014 : 17-18). Ces cooccurrences dépendent toujours de l'empan dans lequel elles sont considérées.

¹¹³ Il est également important de mentionner que toutes ces notions font référence à d'autres méthodes de calculs statistiques. Ce n'est ainsi pas seulement une question de « vocabulaire ».

6.3.4. L'empan

La dernière notion que nous allons définir ici est celle de l'empan. L'empan est la « fenêtre déterminée du texte » (Mayaffre 2014 : 4) dans laquelle nous étudions notre « mot pivot » qui est, dans notre cas, un des quatre verbes modaux. Plus spécifiquement, l'empan représente le nombre d'items qui se trouve avant et/ou après notre mot pivot. Ainsi, si nous voulons, par exemple, rechercher les complémentations verbales du verbe *pouvoir*, *pouvoir* sera le pivot, et l'empan sera le nombre d'items après ce verbe dans lequel nous recherchons les complémentations verbales. L'empan est important, voire « le choix le plus important de la chercheuse » (Evert 2009 : 1221). En effet, le résultat de la recherche n'est pas le même si on étudie ce qui se situe vingt mots ou deux mots avant et/ou après le verbe modal. Pour notre étude, la fenêtre est un empan de trois mots après le modal pour la complémentation verbale et de quatre mots avant le verbe lorsque nous en recherchons les sujets statistiquement associés¹¹⁴. Cet empan a été déterminé après plusieurs tests. Ainsi, nous obtenions par exemple une cooccurrence positive entre *devoir* et *demander* lorsque nous avons un empan plus large mais comme nous pouvons le voir ci-dessous, cette cooccurrence significative n'a pas lieu d'être dans une étude comme la nôtre :

64. Corneille **dût** même **se résoudre à demander** l'aide à trois auteurs (LM, 27.10.2010)

Le verbe *demander* est ici lié à *se résoudre* et non pas au verbe *devoir*. C'est à partir de ce type de résultat que nous avons décidé de resserrer notre empan. Pour les sujets, nous avons également trié plusieurs échantillons avant de déterminer la fenêtre adéquate. Par exemple, pour *devoir*, la plupart des sujets (86%), se trouvent dans les quatre mots qui précèdent le verbe, c'est pourquoi cet empan a été adopté.

6.4. Les calculs et représentations des résultats

Dans notre recherche, nous allons, dans un premier temps, dresser un profil combinatoire pour *pouvoir*, *devoir*, *falloir* et *vouloir*. Le profil combinatoire se définit « comme la structure

¹¹⁴ Notre but est d'étudier les cooccurents en lien direct avec le verbe modal et non pas ceux éloignés de plusieurs phrases, voire de plusieurs paragraphes. Bendinelli (2012), par exemple, considère une fenêtre de 1000 mots pour étudier les auxiliaires et semi-auxiliaires de l'anglais dans les discours présidentiels américains, car son approche est davantage textométrique – donc centrée sur le texte – que se rattachant à la linguistique de corpus – et donc à l'étude de certaines unités lexicales en particulier. Elle étudie ainsi les verbes modaux dans des discours des présidents américains selon le président en question.

schématique du voisinage syntaxique et sémantique d'un mot-pivot telle qu'elle se manifeste dans un vaste corpus » (Blumenthal 2011 : 54). Cette étape nous permet de relever leurs propriétés sémantiques individuelles. Cependant, nous avons aussi pour but de dépeindre leurs ressemblances et divergences. Pour répondre à ces divers objectifs, nous devons ainsi faire appel à plusieurs méthodes de calculs statistiques.

La première que nous présentons ici est l'indice de spécificité (6.4.1.). Ce calcul peut faire émerger les préférences cotextuelles pour chaque verbe modal que nous étudions et donc de dresser leur profil combinatoire. Pour comparer les verbes entre eux, nous ferons appel à l'analyse factorielle des correspondances (désormais AFC) (6.4.2.) et à l'analyse arborée (6.4.3.).

6.4.1. L'indice de spécificité

L'indice de spécificité¹¹⁵ permet de calculer les fréquences relatives et de voir si les cooccurrences que nous étudions sont significatives ou non. Cet indice donne l'occasion d'avoir « des données plus tranchées que la simple distribution des fréquences, car [il] traduit cette distribution en termes de sous-emploi ou de suremploi » (Leblanc et al. 2017 : 141).

Les calculs de spécificité comptent, dans le cas de la cooccurrence, « [...] le nombre de cooccurrences réellement observé avec une estimation de cette fréquence pour le cas hypothétique où les deux éléments (pivot et cooccurent) seraient distribués de manière complètement aléatoire dans le corpus » (Wandel 2020 : 111). Pour expliciter ce calcul, nous recourons à la métaphore d'un sac rempli de X boules, où le sac représente le corpus et X le nombre total de mots du corpus. Les logiciels utilisés pour le calcul de l'indice de spécificité déterminent le nombre attendu d'occurrences des boules A et B parmi X, en prenant pour base leurs fréquences respectives et sur le nombre total de X. Cependant, lorsque plusieurs boules sont effectivement tirées du sac, il est possible que les occurrences de A et B diffèrent de ce qui était estimé initialement. L'indice de spécificité reflète cet écart entre les prévisions probabilistes et les résultats observés, incarnant ainsi la différence entre ce qui était prévu et ce qui se produit réellement lors de ce « tirage » des boules du sac. Ce calcul est tiré du modèle

¹¹⁵ Il existe d'autres méthodes de calculs pour mesurer si les cooccurrences sont significatives ou non, notamment le *log-likelihood*, ou encore le test du *chi-carré* pour ne citer qu'elles. Nous n'avons pas comme prétention d'être une experte en statistiques et donc de connaître tous les avantages et inconvénients de chaque méthode. Cependant, il nous semble important ici d'expliquer le fonctionnement de celle que nous utilisons afin que les résultats soient compréhensibles et transparents dans le chapitre suivant. Nous ne cachons pas le fait que le choix s'est porté sur cet indice également par commodité, car cet indice est celui implanté dans le logiciel TXM que nous avons utilisé pour l'exploration de nos corpus. Pour une explication plus détaillée d'autres méthodes permettant de calculer la cooccurrence, la lectrice peut se référer notamment à la thèse d'Evert (2005).

hypergéométrique qui « est fondé sur la distribution en probabilité du nombre de rencontres de toutes les permutations possibles des formes étudiées dans l'hypothèse d'équiprobabilité » (Née et al. 2017 : 115).

Dans notre travail, nous utilisons comme mesure d'association celui par Lafon (1980), fondée sur ce type de calcul et donc sur ce modèle hypergéométrique. C'est ce calcul qui est implanté dans le logiciel TXM (Heiden et al. 2010), présenté dans la section 6.5.2.1. Pour cet indice, on parle d'exposants et ainsi, « [u]n exposant de valeur 2 exprime une probabilité de l'ordre du centième, 3 du millième, etc. L'absence d'exposant indique que l'usage ne présente pas de caractéristique remarquable. On dira que la forme est banale pour la partie considérée » (Leblanc et al. 2017 : 141). Montrichard (2020) précise que cet « [...] indice de spécificité positif et supérieur à 2 indique que l'unité est surreprésentée par rapport aux autres parties et à la fréquence totale de l'unité dans le corpus » (Montrichard 2020 : 221)¹¹⁶.

Ces données chiffrées peuvent ainsi faire émerger des cooccurrences statistiquement significatives entre unités lexicales. L'étude des cooccurrences permet de révéler un grand nombre d'informations sur les items étudiés car « [l]es répétitions, redites et rafales de mots repérés par l'appareil lexicométrique correspondent dans le discours à des préférences d'emploi sémantiques ou trahissent la récurrence de patrons syntaxiques » (Martinez 2008 : 799-800).

Comme annoncé, ce type de calcul permet ainsi de dresser un profil combinatoire pour chacun de nos verbes. Toutefois, afin de les comparer entre eux, il nous semble pertinent d'utiliser des méthodes statistiques additionnelles.

6.4.2. L'analyse factorielle des correspondances

La méthode de l'analyse factorielle des correspondances (AFC) a été développée par un mathématicien et statisticien, Jean-Paul Benzécri, et exposée dans son ouvrage *L'analyse des données* (1973). C'est donc en premier lieu dans le domaine des mathématiques que cette approche a été mobilisée dans les années 1970 avant d'intégrer les sciences sociales, en particulier la psychologie (Poudat & Landragin 2017 : 103) puis les recherches en linguistique (Mayaffre 2004 : 44-45).

Comme son nom l'indique, l'AFC est une méthode factorielle. Décrites ainsi, ces méthodes

¹¹⁶ Comme pour Pincemin (2012) citée dans la partie 6.3.2., Montrichard (2020) s'insère dans la tradition de la textométrie. Cependant, ses remarques restent valables pour la surreprésentation d'une unité lexicale par rapport à d'autres et non pas seulement selon des parties du corpus.

[...] visent à résumer de manière synthétique des ensembles de données par le biais d'un nombre plus restreint de variables artificielles nouvelles, qu'on appelle des facteurs. Les facteurs sont généralement représentés sous forme de visualisation graphique où les objets à décrire deviennent des points dans un espace. Les méthodes factorielles sont ainsi des outils de synthèse qui offrent un résumé visualisable des contrastes principaux (les plus forts, les plus généraux) dans un tableau de données. (Poudat & Landragin 2017 : 104)

Les AFC sont particulièrement intéressantes pour des tableaux qui comprennent de nombreuses données¹¹⁷ car elles permettent « de représenter globalement, au plus juste les grandes oppositions qui sous-tendent un corpus, d'en repérer les faits saillants en termes de proximités » (Leblanc 2015 : 43-44). Concrètement,

[u]ne AFC vise à situer les partitions du corpus et/ou les variables linguistiques testées, les unes par rapport aux autres, dans un espace multidimensionnel, à partir d'un tableau de contingence à 4 colonnes minimum. Pour la représentation graphique, l'espace multidimensionnel est rapporté à un espace à deux dimensions, les deux axes conservés condensent l'information à même d'opposer les partitions du corpus et/ou les variables linguistiques envisagées. (Bendinelli 2012 : 328)

C'est donc par des points situés sur deux axes – le premier à l'horizontal appelé abscisse, et le second à la verticale, l'ordonnée – que nous pouvons représenter, dans notre recherche, les verbes modaux dans un espace à deux dimensions.

Cette visualisation est complémentaire aux calculs de spécificité car elle donne l'occasion « de jeter un regard plus *macroscopique* sur l'information de base » (Lebart & Salem 1994 : 79). En effet, l'AFC offre une représentation graphique, permettant ainsi une visualisation instantanée des similarités et des disparités les plus significatives entre nos verbes modaux, en fonction des critères sélectionnés. En effet, si l'indice de spécificité permet de mettre en évidence les particularités propres à chaque élément, la comparaison entre ces derniers peut s'avérer plus fastidieuse lors de l'utilisation de cette mesure.

Généralement, la lecture d'une AFC se réalise en opposant la gauche et la droite du graphique, ainsi que le haut et le bas de celui-ci (Veniard & Sitri 2017 : 175-176) car une AFC fait ressortir avant tout ce qui est inattendu, c'est pourquoi ce sont les extrémités du graphique qui doivent être considérées dans un premier temps, c'est-à-dire, l'écart par rapport à la norme qui est représentée par le centre des deux axes (Leblanc 2017 : 218). Ainsi,

¹¹⁷ Les tableaux comportent, par exemples, en entête les divers verbes que nous voulons étudier, dans les lignes, les mots avec lesquels nous voulons étudier leurs cooccurrences, cooccurrences qui sont ainsi représentées dans les diverses cases du tableau. Ainsi si nous voulons étudier les cooccurrences des quatre verbes modaux que nous analysons ici avec tous les verbes qui apparaissent en position de complémentation verbale, nous aurions un tableau de données relativement conséquent, qui peut ainsi être « résumé » à travers l'AFC.

on se concentre d'abord sur les éléments qui s'opposent le plus sur les axes. Ce sont typiquement les éléments les plus éloignés, les plus décentrés qui attirent d'abord l'attention. Les points proches de l'origine de l'axe ne ressortent pas significativement pour observer les contrastes : les observations linguistiques les plus centrées se distribuent ainsi de manière non significative entre les différents textes, tandis que les textes les plus proches de l'origine sont également les plus proches de la norme du corpus. (Poudat & Landragin 2017 : 108)

Poudat et Landragin (2017) parlent de *textes* car ce sont habituellement ces unités-là qui sont étudiées à l'aide de cet outil. En effet,

[l]'AFC est classiquement utilisée par les études linguistiques de corpus pour explorer la structure lexicale d'un corpus textuel. Elle permet ainsi de visualiser dans le même espace de représentation les textes et les mots qui se rapprochent ou qui s'opposent, et de saisir leurs correspondances. (Poudat & Landragin 2017 : 116)

Il n'est donc pas surprenant que cette méthode soit rarement utilisée dans des études se revendiquant plus largement comme faisant partie de la linguistique de corpus. En effet, cette discipline « recourt aux méthodes factorielles avec parcimonie, plus intéressée par l'exploration des collocations et des unités linguistiques en corpus [...] » (Poudat & Landragin 2017 : 122). En cela, nous pouvons constater que notre thèse présente ici un élément innovant dans la manière d'étudier les cooccurrences.

Malgré les nombreux avantages qu'elle représente, il est important de noter que l'AFC « a [...] une fonction purement descriptive : la consultation des résultats est simplement plus aisée » (Lebart & Salem 1994 : 84-85) et elle ne fournit « que des tendances globales qu'il convient d'affiner par la suite, des pistes de recherche, des indices, mais en aucun cas [elle ne doit être considérée] comme une synthèse infaillible du corpus considéré. » (Leblanc 2015 : 45). Il faut ainsi toujours revenir au tableau initial voire au texte pour pouvoir interpréter les résultats visibles sur l'AFC¹¹⁸. En outre, il est également crucial de mentionner que cette représentation en deux dimensions ne permet pas toujours de représenter la totalité des chiffres du tableau. L'AFC n'est ainsi qu'une représentation partielle des données¹¹⁹ d'où l'importance de considérer les AFC uniquement comme un résumé des données, et non pas comme un graphique en totale adéquation avec les informations du corpus. Il est d'ailleurs rare que l'AFC soit utilisée seule dans les recherches et elle sert davantage d'appui à une autre méthode de calcul comme

¹¹⁸ Cette remarque vaut en général pour n'importe quelle donnée quantitative : un chiffre, sans une contextualisation du texte, n'est pas une donnée entièrement interprétable.

¹¹⁹ L'étude de Diwersy et Luxardo (2016) le souligne très bien. L'axe horizontal de l'AFC qu'ils traitent représentent 8,20% des données et l'axe vertical 7,40%. 84% de l'information contenue n'apparaît donc pas dans le graphique analysé.

l'indice de spécificité. En effet, « on combine en général [une recherche basée sur des indices de spécificité] à une AFC, afin de disposer en parallèle d'une image globale des regroupements, qui permet d'apprécier les distances et les relations entre les catégories observées » (Poudat & Landragin 2017 : 176).

Cette méthode n'est pas fréquemment employée dans le cadre des études de cooccurrences¹²⁰. En effet, elle a principalement été utilisée pour des analyses relevant de la textométrie et non pas de la linguistique de corpus. Ainsi, elle a pu démontrer les ressemblances ou les divergences entre différentes parties d'un corpus, comme par exemple, les diverses autrices qui le composent¹²¹. Une AFC de cooccurrences, comme une AFC fondée sur les parties d'un texte, représente « des écarts à l'indépendance sur un graphique à deux axes fondés sur un calcul statistique et probabiliste » (Rossari et al. 2022b : 7). Cependant, contrairement à l'indice de spécificité, les AFC se construisent sur la base des cooccurrences brutes que les items présentent les uns avec les autres. Ainsi, cet « écart à l'indépendance » se calcule non pas de manière proportionnelle par rapport à l'ensemble des formes et de la taille du corpus ou des parties mais relativement aux co-fréquences des formes étudiées avec d'autres formes. Dans notre cas, les AFC pourront par exemple montrer comment *pouvoir* se situe par rapport à certain type de sujets grammaticaux en fonction de comment les autres verbes cooccurrent également avec ces sujets grammaticaux.

Dans notre recherche, nous allons utiliser cette méthode pour étudier en quoi les verbes modaux que nous avons sélectionnés se ressemblent ou divergent les uns par rapport aux autres en fonction de leurs co-fréquences respectives avec les deux paramètres que nous avons sélectionnés, soit les noms en position de sujet ainsi que les verbes à l'infinitif qui se trouvent en position de complémentation verbale.

¹²⁰ Dans le logiciel que nous utilisons pour construire les AFC (Hyperbase, présenté en 6.5.2.2.), il est possible de construire une AFC de cooccurrences généralisées qui se fonde, non pas sur des cooccurrences choisies – par exemple les verbes à l'infinitif qui suivent les modaux que nous étudions – mais dont la méthode « consiste à repérer les rencontres fréquentes ou peu attendues d'une liste de formes, le plus souvent avec elles-mêmes. En d'autres termes, le calcul des cooccurrences généralisées revient à extraire d'un corpus les “items” les plus fréquents ou les mieux répartis (forme graphique ou données lemmatisées en fonction du paramétrage choisi) et de repérer les cooccurrents les plus fréquents de chacune de ces formes » (Née et al. 2017 : 117). Elle s'appuie ainsi principalement sur les cooccurrences des mots les plus fréquents du corpus ou des parties du corpus et leurs cooccurrents les plus fréquents.

¹²¹ Voir par exemple Mayaffre et al. (2019) pour une étude des similarités et des différences entre les différents discours des Présidents de la cinquième République française.

6.4.3. L'analyse arborée

L'analyse arborée est une méthode qui repose sur les mêmes tableaux de données que ceux employés pour les AFC et repose ainsi sur

la distance entre les profils distributionnels des objets considérés, elle propose une représentation graphique sous forme d'arbre permettant d'apprécier les distances entre les points, et de mettre au jour leur nœuds de regroupement ainsi que l'ordre d'apparition de ceux-ci. (Poudat & Landragin 2017 : 135-136)

Plus précisément, ce type de graphique

représente à la fois une classification automatique et un graphe de distances. Les partitions ou les variables sont associées deux par deux, en commençant là où les distances sont les plus petites, jusqu'à avoir regroupé l'ensemble des branches. Le centre topologique de la représentation est indiqué par un petit cercle noir¹²² : le point de raccordement de deux branches est nommé "nœud" : deux (ou plus) branches raccordées sont dites "sœurs". (Bendinelli 2012 : 337)

Ainsi, dans l'analyse arborée, nous parlons de *branche* pour calculer la distance entre les *feuilles*, qui sont représentées dans notre cas par les verbes modaux.

L'analyse arborée peut se lire de diverses manières : on peut se concentrer sur la proximité entre les feuilles, sur les groupes de textes ou de mots qui sont réunies sur une branche, ou alors, comme pour les AFC, sur les oppositions (Barthélémy & Luong 1998 : 7)¹²³.

Cette méthode est, comme pour les AFC, utilisée à l'origine dans le domaine de la textométrie. Gambette et al. (2012) déconseillent l'analyse arborée pour l'étude des distances et rapprochements entre les mots car « il est difficile d'[y] interpréter les nœuds internes, et d'en déduire une façon naturelle d'interpréter la longueur d'une arête située entre deux nœuds internes de l'arbre » (Gambette et al. 2012 : 131). Il a toutefois été noté que l'analyse arborée « est souvent exploitée de manière complémentaire à une AFC » (Poudat & Landragin 2017 : 136) car elle « corrobore les éléments de structure mis au jour avec l'AFC en les précisant » (Mayaffre et al. 2019 : 112). Les distances et les proximités sont davantage lisibles sur ce type de représentation.

L'analyse arborée présente aussi l'avantage de ne plus avoir « à distinguer et à croiser des facteurs, dont chacun n'explique qu'une partie de la variance. Toute l'explication se résume ici en une seule représentation graphique, qui peut prendre deux formes : rectangulaire ou

¹²² Cette représentation est propre au logiciel Hyperbase que nous présentons en 6.5.2.2.

¹²³ Nous adopterons ces différentes interprétations lorsque nous étudierons les verbes modaux via les analyses arborées (7.1.6. et 7.2.5.).

radiale¹²⁴ » (Brunet 2011 : 53). Les distances de la présentation radiale sont plus faciles à interpréter car « elles sont directement proportionnelles à la longueur des parcours » (Brunet 2011 : 54) donc à la distance entre deux mots ou de parties du corpus (par exemple entre les autrices qui composent le corpus). En revanche, la méthode rectangulaire a l'avantage de représenter les données de manière plus condensée encore. C'est cette dernière que nous allons employer car elle permet une meilleure lecture des résultats lorsque ceux-ci ne concernent que peu de paramètres, comme c'est le cas de cette étude se concentrant sur quatre verbes modaux. Rappelons que, tout comme pour l'AFC, « nous ne sommes pas dans une démarche de preuve [avec l'analyse arborée] mais dans une démarche exploratoire. Il s'agit d'accompagner et d'objectiver la lecture, l'analyse et l'interprétation du corpus » (Poudat & Landragin 2017 : 141), c'est pourquoi un retour au texte est toujours nécessaire pour comprendre les graphiques proposés par ces deux méthodes.

6.5. Les outils et corpus employés dans cette thèse

Dans cette partie, nous allons présenter plus précisément les corpus et outils que nous utilisons dans cette recherche. Nous allons donc, dans un premier temps, présenter les corpus dans lesquels nous étudions les verbes *pouvoir*, *devoir*, *falloir* et *vouloir*. Ensuite, nous aborderons les logiciels employés pour analyser ces corpus en lien avec les notions abordées dans les sections précédentes.

6.5.1. Présentation des corpus

Dans ce travail, nous effectuons nos recherches dans deux corpus différents, soit Le Monde et Wikipédia :

I) Le Monde 2010 : presse nationale, 17'895'009 occurrences, désormais LM¹²⁵ :

II) Wikipédia 2019 : encyclopédie participative, 18'715'455 occurrences, désormais WIKI¹²⁶.

Ces deux corpus remplissent les conditions que nous avons présentées dans la section 6.3.1. Premièrement, ils représentent tous deux des textes authentiques et complets – ce sont à chaque

¹²⁴ Cette représentation est propre au logiciel Hyperbase que nous présentons en 6.5.2.2.

¹²⁵ Le corpus LM a été recueilli et construit par Cyrielle Montrichard et nous-même en téléchargeant les articles via la plateforme *Europresse*, accessible avec un identifiant de l'Université de Neuchâtel.

¹²⁶ Nous remercions Helena Bermúdez Sabel qui a conçu le corpus WIKI.

fois les articles dans leur entièreté et non pas des extraits de textes qui figurent dans le corpus. Concernant les conditions de signifiante, ils ont été constitués dans le but d'étudier des faits langagiers dans de gros corpus – dans une vision donc empruntée à la tradition anglo-saxonne. Les articles composant LM ont tous été tirés de l'année 2010¹²⁷ et relatent ainsi les événements de cette période. Ceux qui composent WIKI ont tous été téléchargés le même jour en 2019, représentant ainsi les états de savoir dans différents domaines à un moment précis.

Pour les conditions d'acceptabilité, tous les articles qui composent chacun des corpus comportent les mêmes informations. Pour LM, nous avons par exemple la date et le nom de l'auteur pour chacun des articles. WIKI ne comporte, quant à lui, pas de métadonnées exploitables¹²⁸.

Enfin, tous deux respectent les conditions d'exploitabilité car ils sont homogènes, comportent un nombre suffisant d'occurrences pour être exploités et ne mélangent pas des genres ou champs génériques distincts.

Ces deux corpus représentent respectivement le genre de la presse quotidienne nationale ainsi que le champ générique de l'encyclopédie collaborative¹²⁹ :

¹²⁷ Cette année a été choisie car la Chaire de Linguistique française de l'Université de Neuchâtel au sein de laquelle nous avons effectué notre projet de thèse comptait déjà un corpus de presse régionale de cette même année, soit *L'Est Républicain*. Par souci de comparabilité entre ces deux genres – et donc selon les critères de Pincemin (1999) – nous avons sélectionné des articles de la même période relatant ainsi potentiellement les mêmes événements.

¹²⁸ Cette condition des métadonnées concerne surtout les corpus conçus dans une démarche de l'analyse du discours et donc de la textométrie. En effet, selon cette perspective de recherche, il est important de connaître l'auteur du texte, le sujet abordé, la date de parution, etc. selon le sujet d'étude. Étant donné que nos corpus ont été conçus dans une démarche de linguistique de corpus – et donc pour étudier des unités lexicales dans de grands corpus – la priorité n'a pas été mise sur ces métadonnées qui permettent, par exemple, de nuancer l'usage des verbes modaux selon les thématiques des articles de WIKI. Ce type de métadonnées permettraient de voir si *pouvoir* est utilisé de la même manière dans des articles de géographie ou de littérature – ou encore selon le type d'article de LM – ce qui donnerait l'occasion d'étudier si les verbes modaux apparaissent davantage dans des éditos que des brèves. Pour une réflexion plus approfondie sur l'importance des métadonnées dans les études textométriques, la lectrice peut se référer à la recherche de Fantoli et Vandersmissen (2019) que nous avons notamment mobilisée pour notre étude sur le traitement des athlètes féminines et masculins dans la presse sportive (Aubry & Pamuksaç 2024).

¹²⁹ Nous empruntons ici la terminologie de Malrieu et Rastier (2001) qui proposent une classification en quatre niveaux hiérarchiques pour catégoriser les textes, soit « les discours (ex. juridique vs littéraire vs essayiste vs scientifique), les champs génériques (ex. théâtre, poésie, genres narratifs), les genres proprement dits (ex. comédie, roman "sérieux", roman policier, nouvelles, contes, mémoires et récits de voyage), les sous-genres (ex. roman par lettres). Au niveau inférieur de la classification, nous trouvons les textes d'un même auteur » (Malrieu & Rastier 2001 : 548). Nous nous sommes déjà appuyée sur leur classification pour une recherche portant sur les genres textuels en didactique du FLE (Pamuksaç & Sanvido 2023) dont nous avons adapté le schéma dans la figure 5.

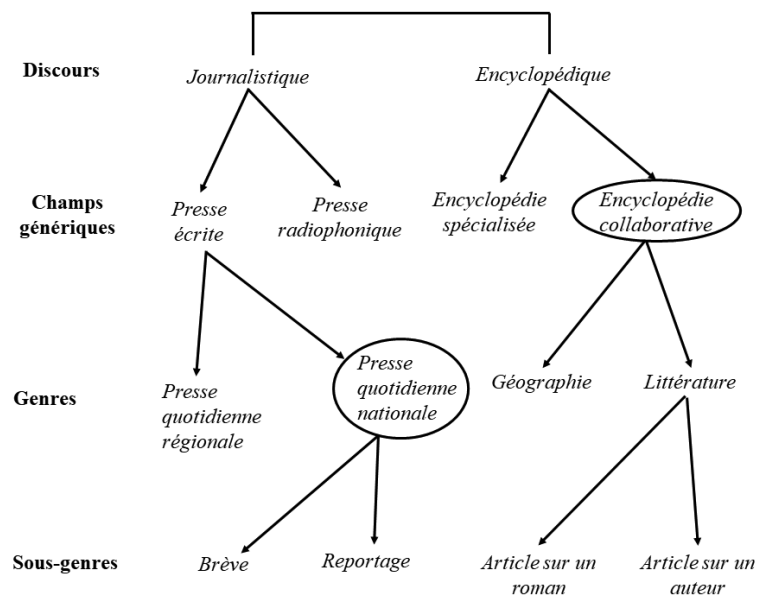


Figure 5 : Niveaux de classification de nos corpus inspirés de Malrieu et Rastier (2001 : 549)

Cette variété de corpus nous permet d’observer si les propriétés que nous relèverons pour les verbes modaux sont stables ou dépendent du texte dans lesquels ces verbes apparaissent. C’est pour cette raison que nous précisons plus haut que notre recherche s’ancre principalement dans la tradition anglo-saxonne de la linguistique de corpus – car nous étudions des items linguistiques en usage – mais dans une moindre mesure également dans une démarche française de la textométrie – car ces items sont comparés dans deux textes différents¹³⁰.

La presse écrite est un genre intéressant pour l’étude des marqueurs modaux. Comme l’explique Hütsch (2020),

ce genre particulier se prête bien à une étude de la modalité par sa caractéristique de combiner différents modes de discours (par exemple, descriptif, narratif ou argumentatif). Les différentes rubriques que l’on trouve dans le genre journalistique peuvent, elles aussi, être le reflet de discours particuliers, comme notamment la rubrique de la une et des faits divers pour un discours souvent purement informatif et factuel, la rubrique des interviews et du courrier des lecteurs pour un discours

¹³⁰ Nous avons fait le choix ici de ne pas fusionner ces deux corpus en un seul macrocorpus pour plusieurs raisons. La première tient du fait que les deux corpus ne figurent pas au même niveau hiérarchique selon la classification de Malrieu et Rastier (2001). De plus, les grands corpus regroupant des genres différents comportent généralement une diversité plus importante et sont majoritairement conçus pour des exploitations plus appliquées. Par exemple, le *Corpus de référence du français contemporain* (CRFC), comptant 310 millions de mots et composé de textes provenant, entre autres, de retranscriptions d’oral spontané, de scripts de films, d’écrits universitaires, etc. « a pour objectif de servir à l’élaboration de dictionnaires, de grammaires et de supports pédagogiques dans le cadre d’une approche à dominante inductive ou “corpus-driven” [...]. Il permettra de prendre en considération certaines questions auxquelles ne permettaient pas de répondre les corpus existants, qu’il s’agisse des variations lexicogrammaticales selon différents genres textuels ou de la phraséologie du français parlé » (Siepmann et al. 2016 : 2). Ces corpus ont aussi pour but de mettre l’accent sur les différences entre ces genres lors d’une recherche.

se rapprochant de l'oral ou la rubrique des sciences et de l'environnement pour un discours spécialisé, c'est-à-dire l'emploi d'une terminologie propre à ces domaines. (Hütsch 2020 : 15)

Ce genre permet ainsi d'avoir accès à un état de la langue prenant en compte beaucoup de variétés, qui « sont autant de regards différents sur un fait d'actualité ou une question de société » (Née & Fleury 2017 : 67). Le genre de la presse a aussi été qualifié de « bon observatoire des pratiques » comme l'ont relevé Dister et al. (2022 : 335). En effet, ce genre de texte est un excellent témoin de la langue telle qu'elle est usitée dans un lieu et à un moment donné car il s'adapte vite aux pratiques langagières¹³¹. La presse peut donc être qualifiée de

“plateforme de lancement” des formules et le lieu privilégié de leur intensification (Krieg-Planque, 2003, 2009), et par conséquent un observatoire de prédilection pour qui s'intéresse à la circulation des mots et des discours et aux rapports de force qui s'exercent dans et par les discours (Foucault, 1971). (Née & Fleury 2017 : 67)

Cependant, il est aussi important de noter que certaines chercheuses voient dans la presse une trop grande diversité langagière pour qu'elle constitue un corpus pertinent à employer dans une recherche linguistique :

En un sens, les corpus constitués de textes de journaux et de données web sont encore moins des corpus prototypiques. Bien que ces corpus soient souvent vastes et relativement faciles à compiler, ils peuvent représenter des registres assez particuliers : par exemple, les articles de journaux sont créés de manière plus délibérée et consciente que beaucoup d'autres textes, ils viennent souvent avec des restrictions linguistiquement arbitraires concernant, par exemple, la longueur des mots ou des caractères, ils ne sont souvent pas écrits par une seule personne, ils peuvent être fortement édités par des éditeurs et des typographes pour des raisons qui peuvent ou non être motivées linguistiquement, etc. Bon nombre de ces conditions peuvent également s'appliquer à certains corpus basés sur le web, bien que les corpus web deviennent de plus en plus des exemples fréquents de l'utilisation de la langue écrite. (Gries & Berez 2017 : 381)

C'est en partie pour ces critiques que nous pensons qu'il est indispensable d'observer les comportements des verbes modaux dans plus d'un corpus afin de pouvoir vérifier si les propriétés que nous relevons sont propres au verbe, ou alors, au type de texte dans lequel il apparaît.

Concernant notre second corpus, à savoir WIKI, il est parfois désigné comme un « corpus de convenance » (Gambette et al. 2021), du fait de la facilité d'accès à ses textes. Cependant, ce

¹³¹ La presse n'intéresse d'ailleurs pas que les linguistes, elle est « un espace privilégié d'observation pour de nombreuses disciplines (sciences de l'information et de la communication, sciences du langage, sciences politiques, sociologie politique, par exemple) » (Née & Fleury 2017 : 67). Fujimura (2005), par exemple, utilise un corpus de presse diachronique afin d'étudier l'évolution de la féminisation des noms de métier en rapport avec des faits sociétaux et des préconisations linguistiques venues d'instances politiques.

corpus est aussi légitime que la presse pour l'étude des phénomènes linguistiques. En effet, WIKI se distingue par sa nature collaborative : l'identification des autrices des articles est ainsi laborieuse voire impossible, ce qui rend difficile la détermination de leur niveau de maîtrise du français, notamment s'il s'agit d'une de leur langue maternelle par exemple. En raison de l'anonymat des autrices mais aussi du champ générique dans lequel ce corpus se situe, WIKI se caractérise par une neutralité de la subjectivité¹³², ce qui rend l'étude des marqueurs modaux particulièrement intéressante dans un contexte où le texte est défini comme objectif.

En somme, les corpus que nous utilisons ici attestent d'un certain état de la langue. Leur diversité nous permet de constater si les propriétés de nos verbes modaux sont stables, selon au moins deux textes différents dans lesquels ils sont usités.

6.5.2. Les logiciels

Dans cette sous-section, nous allons présenter les divers logiciels employés pour cette recherche. Dans le cadre des études se situant dans une perspective de la linguistique de corpus, les outils ont une place primordiale. Selon les investigations à mener, certains logiciels seront plus adaptés que d'autres. Nous n'allons cependant pas présenter les logiciels de manière exhaustive mais uniquement nous attarder sur les fonctionnalités que nous utilisons dans notre recherche.

6.5.2.1. TXM

Les deux corpus que nous avons présentés sont insérés sur la plateforme nommée TXM¹³³ (Heiden et al. 2010) conçue à l'ENS de Lyon. Celle-ci a l'avantage d'être en continuel développement et de jouir ainsi de mises à jour régulières. Nous l'utilisons principalement pour les fonctions suivantes :

¹³² Sur la page d'accueil de Wikipédia (consultée le 04.07.2024) nous pouvons lire : « Ce projet vise à offrir un contenu librement réutilisable, **objectif** et vérifiable, que chacun peut modifier et améliorer. »

¹³³ Pour **TeXtoM**étrie, ce qui démontre bien que le logiciel a été, de prime abord, conçu pour les recherches adoptant une démarche textométrique. Cependant, ses fonctionnalités s'appliquent tout aussi bien à une analyse davantage orientée vers la linguistique de corpus.

- a) Obtenir **les propriétés d'un corpus**, c'est-à-dire les informations concernant la conceptrice du corpus, la taille de celui-ci, les propriétés attribuées aux mots – donc les étiquettes à mobiliser pour les requêtes – etc.

Propriétés de MONDE (CQP ID=MONDE)

Description

montrichardc

25 octobre 2021, 17h24

Statistiques Générales

- Nombre de mots : 17'895'009
- Nombre de propriétés de mot 4 (word, frlemma, frpos, verbe)
- Nombre d'unités de structure 2 (s, text)

Propriétés des unités lexicales (max 100 valeurs)

- frlemma : ", parmi, votre, résolution, de, le, nouveau, année, ,, vous, avoir, décider, changer, opérateur, ., or, il, être, aussi, facile, souscrire, un, abonnement, à, compagnie, assurance, société, téléphone, mobile, ou, fournisseur, accès, Internet, que, difficile, se, en, défaire, pour, acheter, forfait, chez, Bouygues, Telecom, par, exemple, ne, même, pas, nécessaire, déplacer, :, coup, fil, suffire, munir, son, pièce, identité, relevé, bancaire, et, carte, bleu,
- frpos : PUN:cit, PRP, DET:POS, NOM, DET:ART, ADJ, PUN, PRO:PER, VER:pres, VER:pper, VER:infi, SENT, ADV, KON, PRO:REL, NAM,
- verbe : __UNDEF__
- word : ", Parmi, vos, résolutions, de, la, nouvelle, année, ,, vous, avez, décidé, changer, d', opérateur, ., Or, il, est, aussi, facile, souscrire, un, abonnement, à, une, compagnie, assurances, société, téléphone, mobile, ou, fournisseur, accès, Internet, qu', difficile, s', en, défaire, Pour, acheter, forfait, chez, Bouygues, Telecom, par, exemple, n', même, pas, nécessaire, se, déplacer, :, coup, fil, suffit, muni, sa, pièce, identité, relevé, bancaire, et, Carte, bleue,

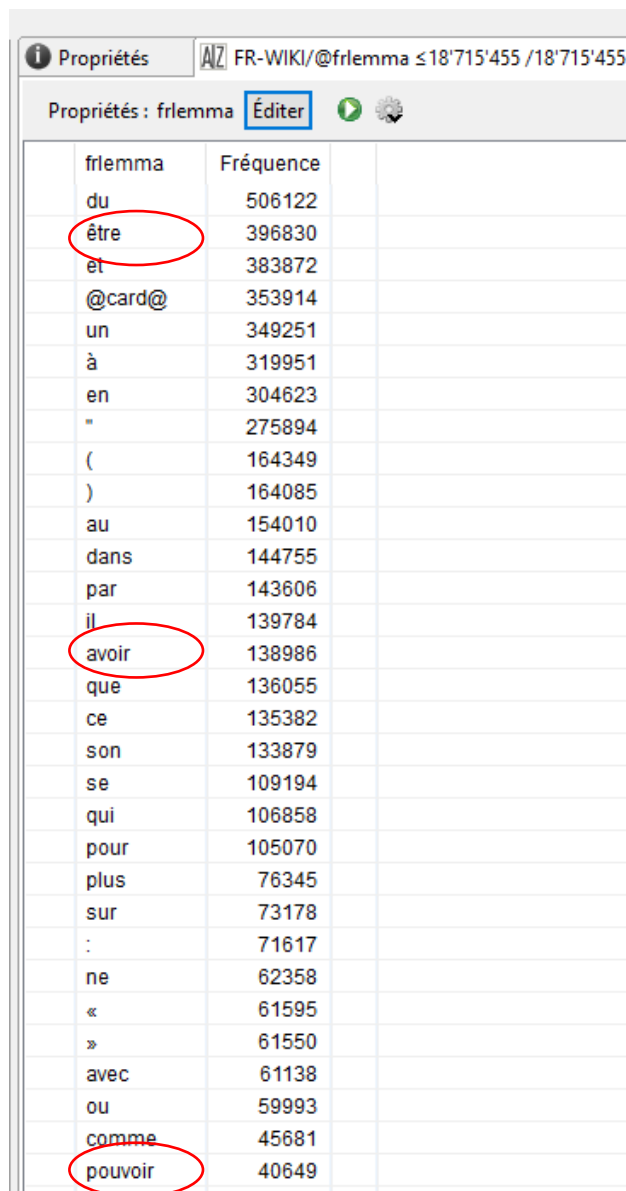
Figure 6 : Capture d'écran de la fonction « propriétés » dans TXM avec le corpus LM.

Sur la figure 6, nous pouvons, par exemple, prendre connaissance du nombre de mots qui composent le corpus (18'669'845) mais aussi les divers *frpos* dans les propriétés des unités lexicales. *Frpos* est une dénomination pour désigner les *parts of speech* avec comme précision la langue, donc *fr* pour français. Cela correspond aux étiquettes morphosyntaxiques¹³⁴ attribuées à des unités lexicales. On peut ainsi y trouver des verbes (VER) avec l'indication du temps ou du mode verbal (VER:pres pour les verbes au présent de l'indicatif par exemple). Ces étiquettes sont importantes pour spécifier la nature grammaticale des termes que nous recherchons. Ainsi, nous pouvons par exemple spécifier que nous recherchons le verbe *pouvoir* suivi d'un verbe à l'infinitif¹³⁵ mais aussi préciser que nous recherchons bien *pouvoir* en tant que verbe et non pas comme nom.

¹³⁴ Les textes entrés dans TXM peuvent être étiquetés et lemmatisés à l'aide de TreeTagger (Schmid 1996).

¹³⁵ Pour cela, nous devons utiliser le langage *Corpus Query Language* (CQL). Pour mener cette requête, cela donnerait : [frlemma = "pouvoir" & frpos = "VER.*"] [][?][frpos = "VER:infi"]. Si

- b) **Le lexique**, qui permet de fournir la liste des formes graphiques du corpus, classées de la plus fréquente à la moins fréquente.



frlemma	Fréquence
du	506122
être	396830
et	383872
@card@	353914
un	349251
à	319951
en	304623
"	275894
(	164349
)	164085
au	154010
dans	144755
par	143606
il	139784
avoir	138986
que	136055
ce	135382
son	133879
se	109194
qui	106858
pour	105070
plus	76345
sur	73178
:	71617
ne	62358
«	61595
»	61550
avec	61138
ou	59993
comme	45681
pouvoir	40649

Figure 7 : Capture d'écran de la fonction « lexique » dans TXM avec le corpus WIKI.

Grâce à la fonction *lexique* nous pouvons, par exemple, observer que le verbe *pouvoir* est le verbe le plus fréquent après les auxiliaires *être* et *avoir* dans le corpus WIKI.

nous « traduisons » cette requête en langage courant cela donnerait quelque chose de type : cherche le verbe (frpos = "VER.") *pouvoir* sous n'importe quelle forme conjuguée (frlemma = "pouvoir") suivi d'un verbe à l'infinitif (frpos = "VER:infi") dans une fenêtre de un à trois mots après *pouvoir* ([]? []? [frpos = "VER:infi"]).

- c) **L'index** qui calcule, par exemple, la fréquence des différentes formes d'un lemme¹³⁶ ou d'une propriété morpho-syntaxique. Nous pouvons ainsi obtenir toutes les formes du verbe *devoir* pour déterminer quel temps ou quelle personne est la plus souvent employée dans un corpus.

Requête 		[frlemma="devoir" & frpos="VER.*"]	
word_frpos	Fréquence		
doit_VER:pres	7468		
devrait_VER:cond	5039		
doivent_VER:pres	2684		
devait_VER:impf	2474		
dû_VER:pper	2305		
devraient_VER:cond	1652		
devra_VER:futu	1114		
devront_VER:futu	765		
devaient_VER:impf	764		
devoir_VER:infi	628		
devons_VER:pres	575		
due_VER:pper	417		
dois_VER:pres	206		
dues_VER:pper	172		
devant_VER:ppre	150		
dus_VER:pper	127		
devrions_VER:cond	66		
devais_VER:impf	64		
devez_VER:pres	63		
dut_VER:simp	50		
Doit_VER:pres	45		
Devant_VER:ppre	43		
devions_VER:impf	37		
devrais_VER:cond	37		

Figure 8 : Capture d'écran de la fonction « index » dans TXM avec le corpus LM.

Dans la figure 8, nous pouvons observer que *devoir* est le plus souvent conjugué à la troisième personne du singulier à l'indicatif présent, puis à la même personne mais au conditionnel présent et ensuite à la troisième personne du pluriel à l'indicatif présent.

- d) **Le concordancier** qui permet de lire toutes les occurrences de notre requête sous forme de *Key Word In Context* (désormais KWIC). Ceci nous permet d'opérer un retour au texte, élément primordial lorsque nous faisons de la linguistique avec des outils

¹³⁶ La forme lemmatisée d'un mot « revient à transformer un texte en groupant sous une forme canonique (telle qu'on la trouve dans les dictionnaires) tous les substantifs, les adjectifs, les verbes ainsi que les formes élidées (*j'*, *t'*, *l'*, etc.). » (Née et al. 2017 : 106). Pour les verbes, il s'agit par exemple de prendre en considération toutes les formes de ceux-ci, donc tous temps verbaux et personnes confondues.

statistiques (Leblanc 2015 : 53). Le concordancier n'est pas un outil qui est né avec l'émergence de l'informatique. Comme le rappellent Leblanc et al. (2017),

[l]e principe en est d'ailleurs fort ancien et ne doit rien à l'analyse statistique des données textuelles, qui n'a fait que "l'implémenter". Dès le xii^e siècle, saint Chef effectuait des concordances de la Bible et on ne s'étonnera pas que les premières concordances aient concerné les textes religieux puisqu'il s'agissait de déterminer le sens d'un mot, voire tous les sens d'un mot, en le replongeant dans toutes ses réalisations et en le replaçant dans son contexte immédiat. (Leblanc et al. 2017 : 143)

Dans TXM, cet outil se présente ainsi, avec le co(n)texte gauche, le pivot et le co(n)texte droit divisé en trois colonnes :

Contexte gauche	Pivot	Contexte droit
l'auteur des Fausses Confidences atteint la soixantaine. Courte et tardive	veut	dire, ici, que Marivaux y livre une sorte de quintessence
appellation actuelle, CulturesFrance ? « Il dit bien ce qu'il	veut	dire, insiste le sénateur Louis Duvernois. Sauf qu'il donne
, l'homme qui m'a appris ce qu'être un homme	veut	dire, l'homme qui a rendu à jamais impossible que je
ignorants qui n'ont aucune espèce d'idée de ce que philosophie	veut	dire, la répétition pavlovienne de la scie - sonnait comme un
, Par-dessus bord est devenue Tori no tobu takasa, ce qui	veut	dire, peu ou prou, « la hauteur à laquelle volent
une croissance à 1 % -1, 5 % ». Cela	veut	dire, poursuit-il, que la France ne peut plus continuer
on referme Choir, on ne sait plus trop ce que roman	veut	dire, tant il a déjoué nos attentes et suscité de la
A condition de ne pas reproduire les mêmes erreurs. Ce qui	veut	dire, tout d'abord, éviter la multiplication des candidatures dissidentes
soient mêlés dans le livre », raconte Bureau. Ce qui	veut	dire : je suis un journaliste, pas un artiste. Mais
affirmé que j'étais, mais attends un peu -, je	veux	dire : Mon chéri chéri, tu as tué le passé.
défaillante de la principale centrale thermoélectrique du pays. « Ce qui	veut	dire : plus de frigos et plus de ventilateur », illustre
, depuis presque trente ans, est un temps désorienté. Je	veux	dire : un temps qui ne propose à sa propre jeunesse,
la réforme des retraites. Concrètement, qu'est-ce que cela	veut	dire ? - Il est clair que les mesures générales que l'
et médecins. Faire l'amour, qu'est-ce que cela	veut	dire ? A ce sujet, les petits n'ont pas la
quoi bon \ " ? Etre libre qu'est-ce que cela	veut	dire ? Etre libre toute seule ? Ma mère n'était plus
retraites... Un Grenelle des retraites, qu'est-ce que ça	veut	dire ? Qu'on fait une vaste concertation, qu'on décide
définitive, nous serions tous éternels, vous voyez ce que je	veux	dire ? Tous immortels. Alors que nous sommes terriblement fragiles et
ne sont pas des juifs, si vous voyez ce que je	veux	dire. » Un sentiment que vient vigoureusement appuyer l'un de
il a gravi tous les échelons, il sait ce que lutter	veut	dire. Depuis onze ans qu'il dirige la CGT, il
ce siècle se méfie de ce que \ " créer \ "	veut	dire. Désormais, on baptise \ " créateurs \ " des
de Libération, Blandine Grosjean, qui sait ce que l'époque	veut	dire. Elle revendique le titre de rédactrice en chef de Paris
la chaîne d'habillement Topshop, sait ce que gérer un groupe	veut	dire. Et très clairement, ce qu'il a vu de
simplement cessé de réfléchir à ce que les droits de l'homme	veulent	dire. Il est pourtant devenu clair que le triomphe d'un
n'est pas une aube, si vous voyez ce que je	veux	dire. On est passé directement à la conclusion. A propos
pas le mot qui compte, c'est ce que la personne	veut	dire. Par exemple, « merde » ou « putain »
des CCN de Tours et de Rennes, savent ce que durer	veut	dire. Résister aussi. Larrieu a décidé de quitter Tours il
« gouvernement économique », mais c'est bien ce que cela	veut	dire. Un petit pas, au moins sémantique, est accompli
Et toujours cette élégance plastique sans faille qui sait ce que beauté	veut	dire. Un vieil homme, assis sur une chaise, de
sonne turc, comme Mehmet par exemple... et Mehmet, ça	veut	dire... Ah ! Vindiou ! Inch'Allah ! C'est sûr qu'
dont tout le sens est de nous dire ce qu'être Français	veut	dire... Et je ne parle même pas des suiveurs qui,
verbe « encadrer » (...) est une formulation militaire qui	veut	dire « assister » ». "

Figure 9 : Capture d'écran de la fonction « concordances » dans TXM avec le corpus WIKI.

Dans la figure 9, nous pouvons observer le cotexte du verbe *vouloir* au présent. Nous pouvons trier le co(n)texte gauche, le pivot ou le co(n)texte droit par ordre alphabétique. Ici nous avons choisi de trier le contexte droit par ordre alphabétique afin de retrouver toutes les occurrences de *vouloir* au présent avec *dire*.

- e) **Les cooccurrences** qui permettent de calculer l'indice de spécificité (Lafon 1980) entre plusieurs items dans un empan choisi.

Cooccurrent	Fréquence	CoFréquence	Indice	Distance moyenne
KON_que	102952	475	250	1.1
VER:infi_noter	862	112	219	.3
ADV_pas	35341	227	148	.2
ADV_donc	10051	140	136	.0
VER:infi_attendre	723	441	127	.1
VER:infi_ajouter	658	192	127	.3
VER:infi_confondre	254	56	124	1.1
KON_pour	8066	101	94	1.9
VER:infi_distinguer	675	50	85	.3
VER:infi_faire	8133	81	68	.4
VER:infi_considerer	573	40	67	.5
VER:infi_oublier	241	32	63	1.1
VER:infi_preciser	226	31	62	.3
VER:infi_souligner	166	29	61	.5
VER:infi_signaler	141	28	61	.3
VER:infi_prendre	2083	49	59	.3
VER:infi_savoir	1114	39	54	.2
VER:infi_rappeler	223	25	47	.5
ADV_aussi	21737	94	47	.2
VER:infi_chercher	434	28	46	.6
VER:infi_mentionner	111	20	42	.6
VER:infi_rajouter	70	18	41	.1

Figure 10 : Capture d'écran de la fonction « cooccurrences » dans TXM avec le corpus WIKI

Dans la figure 10 apparaissent les cooccurrents du lemme *falloir* qui sont spécifiques dans un empan de trois mots après le pivot¹³⁷. Dans la colonne ‘cooccurrents’, nous pouvons voir de quel cooccurrent il s’agit. Nous avons choisi ici d’avoir comme informations dans les résultats les étiquettes morphosyntaxiques ainsi que la forme lemmatisée de l’unité lexicale en cooccurrence spécifique. Dans la seconde colonne, ‘fréquence’, nous pouvons prendre connaissance de la fréquence dudit cooccurrent. Pour la colonne ‘cofréquence’, le chiffre qui y apparaît correspond au nombre de fois où le pivot et le cooccurrent en question apparaissent ensemble dans cet empan. L’avant-dernière colonne présente l’indice de spécificité qui, nous le rappelons, est spécifique à partir de 2 selon l’indice de Lafon (1980) implémenté dans TXM.

¹³⁷ TXM considère le premier mot avant ou après le pivot comme étant à une distance de 0, c’est pour ça que nous avons fixé ici un contexte de 2 à droite du pivot pour capturer les trois mots qui viennent après *falloir*.

Enfin, la dernière colonne, ‘distance moyenne’, indique, comme son nom l’indique, à combien de mots en moyenne se trouve généralement le cooccurent du pivot.

6.5.2.2. *Hyperbase*

Hyperbase (Brunet 2011) a été créé et développé par Etienne Brunet en 1967 et depuis 2013, il est maintenu à jour par Laurent Vanni. Il est distribué par le CNRS et l’Université Côte d’Azur. Comme TXM, il a été initialement conçu pour les études textométriques, mais nous l’employons dans une optique de la linguistique de corpus anglo-saxonne¹³⁸.

TXM peut fournir des AFC, non pas de cooccurrences « choisies », mais seulement pour les cooccurrences généralisées (6.4.2). Nous aurions pu ainsi produire des AFC qui montrent la distribution de différents éléments par rapport aux partitions d’un corpus. Etant donné que les corpus que nous utilisons dans cette recherche n’ont pas été conçus pour être exploités dans une optique d’analyse du discours, ils ne sont pas partitionnables à l’aide de métadonnées, comme nous l’explicitons en 6.5.1. Si tel avait été le cas, nous aurions pu étudier, par exemple, l’usage des verbes modaux selon des thématiques dans WIKI ou encore selon le type d’article dans LM. Dans Hyperbase, nous pouvons entrer nos propres données sous forme de tableau de contingences¹³⁹ et ainsi étudier les cooccurrences entre deux termes plutôt que leur distribution dans notre corpus. Nous utilisons, pour ce faire, la colonne co-fréquence de la fonction ‘cooccurrences’ de TXM. Ce tableau permet tant de créer des AFC que des analyses arborées dans Hyperbase.

Synthèse sur la méthodologie

Dans cette partie, nous avons présenté les outils méthodologiques employés dans cette recherche afin d’étudier les propriétés sémantiques individuelles des quatre verbes modaux *pouvoir*, *devoir*, *falloir* et *vouloir* à travers leurs profils combinatoires, mais également leurs similarités et disparités avec l’AFC et l’analyse arborée. Dans la littérature à leur sujet, nous avons soulevé le fait que des chercheuses se focalisent sur leur cotexte afin de déterminer les

¹³⁸ Pour des utilisations de l’ordre de l’analyse du discours de ce logiciel, se référer notamment à Brunet (2012) pour les cooccurrences et à la thèse de Kastberg-Sjöblom (2002) pour un panorama plus détaillé des autres fonctionnalités.

¹³⁹ Par tableau de contingences, nous entendons un tableau qui comporte donc les cooccurrences brutes entre les verbes modaux que nous étudions et les différents paramètres pris en compte pour notre recherche.

valeurs modales et/ou les effets de sens de ces verbes, sans recourir toutefois à des méthodes qui permettent d'extraire et d'analyser cet entourage de manière automatique.

Pour notre recherche, nous avons ainsi opté pour les outils de la linguistique de corpus et de la textométrie. Dans un premier temps, nous utiliserons le calcul de l'indice de spécificité pour calculer les cooccurrences significatives entre les modaux et deux éléments de leur valence, soit le sujet grammatical sous forme de nom ainsi que le complément d'objet représenté par un verbe à l'infinitif. Plus un indice est grand, plus l'attraction est forte. Ce type de calcul permet ainsi de représenter le profil combinatoire des verbes étudiés ici. Mais, afin de comparer au mieux les verbes entre eux, nous faisons également appel à des calculs et représentations graphiques qui permettent de mettre en exergue ces similarités ou divergences, plus précisément l'analyse factorielle des correspondances et l'analyse arborée.

Il est rare de trouver une plateforme ou un logiciel qui nous fournisse tous les outils nécessaires pour mener à bien une étude. C'est pourquoi nous nous appuyons sur les fonctionnalités de plusieurs logiciels pour répondre à nos questions de recherche. Nous nous servons ainsi des logiciels TXM et Hyperbase afin de pouvoir tester toutes les hypothèses émises dans le chapitre 5.

TXM et Hyperbase sont des programmes utilisés surtout pour des analyses textométriques. Cependant, il est également important de rappeler que, étant donné que nous analysons les verbes *pouvoir*, *devoir*, *falloir* et *vouloir* dans deux types de textes distincts, notre approche n'est pas en totale opposition avec l'approche textométrique. En effet, nous ne pouvons pas ignorer le type de texte dans lequel ces verbes sont employés, et les propriétés que nous identifions sont toujours spécifiques au corpus dans lequel elles ont été relevées, comme le rappelle, entre autres, Blumenthal (2008) : « Nous ne cessons de souligner que la validité des résultats de cette méthode dépend du corpus choisi. Car le style de “la vie des mots”, pour reprendre le titre d'un livre célèbre de sémantique française, diverge souvent fortement d'un type de textes à l'autre » (Blumenthal 2008 : 35).

Cette combinaison d'outils, tout comme la diversité des corpus étudiés, nous permet ainsi d'étudier les verbes modaux sous plusieurs angles afin d'étayer au mieux les analyses à leur propos.

7. Résultats et analyses

« You shall know a word by the company it keeps » (Firth 1957 : 11) est le principe de base sur lequel nous nous reposons afin de connaître les propriétés sémantiques ainsi que les ressemblances et les divergences des quatre verbes modaux que nous étudions ici. Ce co(n)texte privilégié sera étudié à partir de deux éléments obligatoires pour les verbes modaux soit *le sujet grammatical* ainsi que le *complément d'objet*. Pour le premier élément, nous nous attardons sur les noms qui apparaissent en position de sujet et pour le second sur les verbes à l'infinitif qui sont modalisés par *pouvoir*, *devoir*, *falloir* et *vouloir*. Ces deux éléments ont été sélectionnés pour notre étude car ils sont facilement récupérables automatiquement à l'aide des outils dont nous disposons mais aussi parce qu'ils sont riches sémantiquement, et offrent ainsi une couche de sens supplémentaire aux énoncés. Nous détaillerons ces choix respectivement en 7.1. et 7.2. Dans un premier temps, nous décrirons ces verbes sous leur forme lemmatisée puis nous varierons les temps verbaux. Il semble important de diversifier ces paramètres, car les temps verbaux ont énormément d'influence sur le verbe et son interprétation :

Dans la phrase, les formes verbales apportent d'abord des informations sur le temps : elles permettent la localisation temporelle de la situation : elles apportent également des informations sur l'aspect, c'est-à-dire sur le déroulement de la situation : enfin, elles apportent des informations dites *modales* qui caractérisent l'attitude du locuteur par rapport à la réalisation de la situation. Toutes ces informations apportées par la conjugaison verbale interagissent avec les propriétés sémantiques du verbe et avec celles du reste de la phrase (sujet, compléments et ajouts éventuels). (Vet et al. 2021 : 1227)

Il est également important de préciser que des éléments externes peuvent amener une modification quant à l'interprétation du temps verbal. Ainsi,

L'interprétation des temps verbaux dépend souvent du reste de la phrase et du contexte. Par exemple, sans autre contexte, la phrase [*Je travaille*] est comprise comme décrivant une situation simultanée à l'énonciation. Avec l'ajout de *demain*, l'interprétation change. [*Demain, je travaille*] parle d'une situation postérieure à l'énonciation. Avec *ce matin*, la situation est située avant l'énonciation [*Ce matin, j'arrive et qu'est-ce que je vois ?*] (Vet & de Swart 2021 : 1240)

En plus de la forme lemmatisée, nous étudierons ainsi les verbes modaux au présent de l'indicatif et au conditionnel présent. Le premier temps a été choisi car sa polysémie lui confère non seulement la possibilité d'indiquer « les trois époques différentes de l'actuel, du passé et du futur, qui, théoriquement, s'excluent l'une l'autre » (Touratier 1996 : 74). Son interprétation dépend ainsi d'indicateurs temporels comme susmentionnés mais également « un présent dit

gnomique ou permanent qui apparaît notamment dans les définitions [...] les proverbes [...] les faits scientifiques ou naturels, appelés souvent faits d'expérience [...] » (Touratier 1996 : 77) à tel point que « [c]ertains grammairiens considèrent que le présent est *atemporel*, ou *omnitemporel* dans les phrases *génériques*, ou de vérité générale » (Vet & de Swart 2021 : 1242). Ce type d'usage pourrait ainsi renforcer des valeurs descriptives dans l'interprétation des verbes modaux. Notons également que « le présent peut exprimer une obligation (le présent déontique) » (Stage 2003 : 203) et donc véhiculer une valeur modale déontique.

Quant au conditionnel¹⁴⁰, par sa nature, il exprime une incertitude et permet ainsi de véhiculer, dans une acception large des marqueurs modaux, la valeur modale épistémique. C'est le cas notamment du conditionnel journalistique :

[Le conditionnel] est fréquemment utilisé pour indiquer que l'information donnée dans la phrase a été acquise par ouï-dire. On parle parfois de *conditionnel journalistique*. Du point de vue temporel, la phrase [*Le président de la République serait à New York en ce moment.*] a la valeur d'un présent : elle signifie 'Le président de la République est à New York', mais le conditionnel exprime une certaine réserve du locuteur à l'égard de ce contenu : non seulement l'information est présentée comme non vérifiée, mais comme provenant d'un autre locuteur, situé à un autre repère, antérieur à l'énonciation. Il s'agit donc d'un emploi modal. (Vet & de Swart 2021 : 1254)

Le conditionnel permet également, comme l'imparfait, de « créer une certaine distance entre le locuteur et la réalisation de ces situations : même si elles peuvent être réalisées, le locuteur les présente comme peu probables » (Vet & de Swart 2021 : 1254). Le conditionnel permet d'« atténuer l'expression du souhait ou du désir par le verbe. [...] La fonction du conditionnel est de créer une certaine réserve » (Vet & de Swart 2021 : 1255) comme nous l'avons déjà relevé pour *vouloir* lorsqu'il est conjugué à l'imparfait (4.5.3.).

Rappelons que la variation des effets de sens et des valeurs modales selon les temps verbaux est considérée depuis longtemps par les chercheuses (Flaux & Lagae 2014 : 4). Faire varier ces temps et modes verbaux nous permettra de voir si les propriétés que nous relevons sont stables, peu importe la forme du verbe, ou si celle-ci a une influence sur les valeurs modales et effets de sens véhiculés.

Nous établirons des profils combinatoires pour chacun de ces verbes en analysant leurs associations avec les sujets nominaux et leurs compléments verbaux sous forme de verbe à

¹⁴⁰ Vet et de Swart (2021) proposent de considérer le conditionnel comme « un temps de l'indicatif, interprété comme un futur du passé (*J'ai décidé que je viendrais demain.*). Le conditionnel dit *passé* décrit l'état résultant d'une situation dans le futur par rapport à un passé (*Il m'avait dit qu'il aurait fini pour midi.*) » (Vet & de Swart 2021 : 1253). Nous ne proposons ici pas une réflexion plus approfondie sur cette question car cela ne représente pas le cœur de notre thèse mais il semble important de le signaler ici si des effets de sens de cet ordre devait apparaître dans nos exemples.

l'infinitif. Ces résultats seront examinés à plusieurs niveaux : tout d'abord, à travers leur répartition dans des catégorisations sémantiques de noms et de verbes. Ces catégorisations offrent une vue d'ensemble initiale des résultats et proposent « une grille d'analyse qui permet d'appréhender les éléments à classer et d'identifier leurs propriétés distinctives » (Haas et al. 2022 : 53). Ces catégorisations seront établies sur la base des fréquences brutes des sujets et des compléments verbaux qui sont significativement associés aux modaux.

Ensuite, une seconde phase d'analyse avec l'indice de spécificité sera amorcée. Nous avons choisi d'analyser les 30 cooccurrences en position de sujet ou de complément verbal avec l'indice le plus élevé. Ce choix est motivé par la grande quantité de résultats ainsi que par la pertinence de se concentrer sur des indices de spécificité élevés, illustrant les usages les plus saillants des modaux. À partir de ces indices, nous examinerons les modaux tels qu'ils sont employés avec les noms en position de sujet et les verbes en position de complément verbal les plus significatifs. Cette partie de l'analyse visera à illustrer quelques cas spécifiques de ces verbes sélectionnés à partir d'un échantillon¹⁴¹, sans prétendre à une illustration exhaustive de tous les usages possibles. Cependant, en se basant sur les indices les plus élevés, les exemples choisis représentent parmi les usages les plus significatifs parmi ceux relevés afin d'illustrer le fonctionnement et les caractéristiques sémantiques des verbes modaux.

En résumé, les répartitions initiales permettent d'obtenir un aperçu du profil combinatoire des modaux et de visualiser leurs similitudes et disparités à travers la distribution des catégories sémantiques représentées dans les deux paramètres étudiés. En examinant ensuite de plus près les catégories sémantiques avec lesquelles les verbes modaux s'associent de manière importante, nous pourrions observer leurs usages effectifs ainsi que les effets de sens et les valeurs modales qu'ils véhiculent de manière significative.

Nous comparerons ensuite tous ces verbes à l'aide de l'AFC et de l'analyse arborée pour mettre en évidence les similarités ou les divergences entre eux.

¹⁴¹ Il arrive que le nombre d'exemples soit relativement haut pour certaines cooccurrences. Nous avons ainsi sélectionné au hasard 50 exemples à chaque fois afin de pouvoir illustrer les usages que nous analysons dans ce chapitre. Si les 50 exemples montraient des effets de sens ou valeurs modales trop disparates, nous avons analysé 50 exemples de plus afin de corroborer nos observations du premier échantillon.

7.1. Les noms en position sujet

Un des éléments entourant le modal – et les verbes en général en français – est le sujet grammatical¹⁴². Celui-ci a d'ailleurs une certaine importance pour l'interprétation sémantique des verbes comme le rappelle Girard (2004) avec son exemple du verbe *chanter* :

Chaque fois, la représentation que nous nous faisons du type de son entendu dépend du sujet de l'énoncé : “la Callas”, “mon frère”, “les oiseaux”, etc. : et il y a peu de ressemblance entre ce que produit la Callas, et ce que produit mon frère. Ceci signifie donc que le “sujet”, par sa spécificité, fait émerger un sens particulier de chanter. Je ne dis pas qu'il complémente son sens car la notion “chanter” si elle se distingue d'une autre notion est définissable en elle-même. Je ne dis pas non plus, comme le suggère Adamczewski, que le sujet “limite” la notion, car cela signifierait que dès que le verbe est utilisé dans un énoncé, son sens s'en trouve réduit. Le point important à retenir c'est que le sujet est partie prenante dans l'interprétation que je fais de la notion “chanter” : en ce sens, il joue un rôle sémantique par rapport au verbe, et c'est ce rôle sémantique qui est fondamental, car il le distingue des autres éléments qui sont des “compléments”. (Girard 2004 : 45)

Cette citation rappelle l'importance du co(n)texte pour interpréter le sens des unités lexicales, importance déjà soulevée par Recanati (2004) :

Le mot ‘couper’ n'est pas ambigu, dit Searle, pourtant il contribue de manière très différente aux conditions de vérité de l'énoncé dans ‘Bill a coupé l'herbe’ et ‘Sally a coupé le gâteau’. Cela s'explique par le fait que les présupposés contextuels jouent un rôle dans la détermination des conditions de satisfaction, et différents présupposés contextuels sous-tendent l'utilisation de ‘couper’ en relation avec l'herbe et les gâteaux respectivement. Nous supposons (révocablement) que l'herbe est coupée d'une certaine manière, et les gâteaux d'une autre manière. (Recanati 2004 : 90-91)

En ce sens, nous ne pouvons nier que le sujet grammatical est important pour l'analyse de l'effet de sens d'un verbe – qu'il soit modal ou non – car « [p]ar son référent, le sujet donne [...] un contenu tangible à la notion verbale. On peut même aller jusqu'à dire que le référent du sujet et le référent du verbe sont une seule et même entité » (Girard 2003 : 39). Les exemples 68 et 69 illustrent bien cette citation :

65. Paul ne *comprend* pas la question.

66. L'offre *comprend* le petit-déjeuner.

¹⁴² Par sujet grammatical, nous entendons l'élément qui permet de savoir à quelle personne conjuguer le verbe qu'il régit.

Ainsi, dans 65, le sujet de type *humain* active l'effet de sens *cognitif* du verbe *comprendre* alors que le sujet *objet* de 66 active celui d'*inclure*. Comme Fournier et Fuchs (1998), nous admettons ainsi que nous pouvons « [p]arler de recatégorisation du verbe sous l'effet du trait sémantique (animé ou inanimé) associé au sujet, [et donc] voir dans ce trait une cause de la valeur du verbe [...] » (Fournier & Fuchs 1998 : 67).

Pourtant, pour les verbes modaux, une certaine distinction doit s'opérer. En effet, ces verbes n'ont pas une grande influence sur leur sujet, étant donné qu'ils servent surtout à modaliser le verbe qui suit. C'est cette complémentation verbale sous forme de verbe à l'infinitif qui dicte le type de sujet qui précède le modal. En cela, nous pouvons rappeler les réflexions d'Abeillé et Koenig (2021a), déjà présentées dans le chapitre 4 de ce travail :

Les linguistes distinguent deux types de verbes avec infinitif : ceux comme *vouloir*, souvent dits à *contrôle*, et ceux comme *sembler*, souvent dits à *montée*. Nous les appelons respectivement verbes à *partage* et verbes à *héritage* : dans le premier cas, le sujet de l'infinitif est sélectionné à la fois par le verbe principal et l'infinitif, qui le *partagent* : dans le second cas, il est sélectionné uniquement par l'infinitif et *hérité* par le verbe principal. (Abeillé & Koenig 2021a : 169)

Il y aurait ainsi une importance différente du sujet selon que celui-ci est explicitement sélectionné par le verbe conjugué ou par le verbe à l'infinitif qui suit :

Ainsi, avec un verbe comme *vouloir*, le sujet [...] est interprété comme un agent (celui qui veut). Ce n'est pas le cas avec un verbe comme *sembler* : [dans *Jean semble dormir*,] Jean n'est pas interprété comme 'celui qui semble', mais simplement comme celui qui dort. Du point de vue sémantique, l'argument de *sembler* est l'ensemble sujet + infinitif, c'est-à-dire 'Jean dort'. (Abeillé & Koenig 2021a : 169)

Selon cette distinction, dans notre paradigme des verbes, seul *vouloir* aurait ainsi une influence sur le choix du sujet (plutôt des êtres *animés* si nous admettons son NS d'*intention* ou de *volonté*) alors que *pouvoir* et *devoir* sont uniquement là pour modaliser le verbe qui suit. Il convient de noter, comme mentionné dans le chapitre consacré au verbe *vouloir*, que ce soit dans la section 4.5.1. traitant de son intégration parmi les verbes modaux ou dans la section 4.5.6. concernant ses usages grammaticalisés, que ce verbe ne se distingue finalement pas significativement des verbes *pouvoir* et *devoir*. En effet, « [d]e façon générale, dès qu'un verbe à partage peut être employé avec un sujet non humain, non agentif, et dans un emploi statique, il tend à avoir des propriétés d'un verbe à héritage » (Abeillé & Koenig 2021a : 171). Étant donné que ce type d'emploi n'est pas négligeable pour *vouloir*, il semble pertinent de pouvoir le comparer aux deux autres verbes à sujet que nous étudions ici. De plus, notons que tous les verbes qui demandent un infinitif – et donc pas seulement les modaux – « identifient le sujet

implicite de l'infinitif avec leur propre sujet » (Caudal 2021 : 1325), ce qui aligne finalement *pouvoir*, *devoir* et *vouloir* et montre ainsi que la distinction entre les verbes à partage et les verbes à héritage n'est pas toujours pertinente, étant donné que des caractéristiques syntaxiques sont partagées entre ces verbes.

En plus de cela, rappelons que des linguistes ont concédé à *pouvoir* et *devoir* des interprétations sémantiques selon le type de sujet qui les régit (4.2.4. pour *pouvoir* et 4.3.6. pour *devoir*). Ces observations peuvent sembler contradictoires à première vue. Bien que les verbes modaux tels que *pouvoir* et *devoir* ne soient pas strictement sélectifs quant à leur sujet, le type de sujet utilisé influence dans une certaine mesure leur interprétation sémantique. Cette influence, bien que subtile, justifie l'examen de la proximité des contextes d'utilisation préférés de ces trois verbes pour déterminer les différences sémantiques entre *pouvoir*, *devoir* et *vouloir*. En partant de ce constat et compte tenu des approches quantitatives adoptées dans cette thèse, il est pertinent de s'interroger sur la proximité des contextes d'utilisation préférés de ces trois verbes afin de déterminer si des différences s'opèrent entre *pouvoir*, *devoir* et *vouloir*. Les deux premiers verbes n'imposent pas de contraintes strictes sur le choix du sujet, tandis que *vouloir*, en toute logique, devrait sélectionner principalement des sujets animés. Cependant, comme souligné, nous savons déjà que ce n'est pas toujours le cas. Cette flexibilité dans l'utilisation de *vouloir* permet de justifier son inclusion dans une comparaison avec *pouvoir* et *devoir*, et démontre que la distinction entre verbes à partage et verbes à héritage doit être réévaluée à la lumière des données d'usage réelles.

Pour conclure, bien que *vouloir* soit traditionnellement classé comme un verbe à partage, son usage avec des sujets non-humains montre qu'il présente des propriétés communes avec *pouvoir* et *devoir*, des verbes considérés comme à héritage. Cette fluidité entre les catégories théoriques met en lumière l'importance de considérer les nuances de l'usage dans notre analyse. Ainsi, comparer *vouloir* à *pouvoir* et *devoir* est non seulement justifié, mais nécessaire pour une compréhension plus nuancée des verbes modaux. Les points communs semblent ainsi plus importants que les divergences à leur propos.

Nos questions de recherche pour cette partie sont ainsi les suivantes :

- I) Si *pouvoir* et *devoir* ne sélectionnent, *a priori*, pas leur sujet de manière aussi importante que *vouloir*, mais que c'est le verbe qu'ils modalisent qui le choisit, peut-on quand même noter une préférence pour un certain type de sujet plutôt qu'un autre ?
- II) Si préférence avérée il y a pour certains types de sujets, ces préférences sont-elles associées à des effets de sens ou valeurs modales plutôt que d'autres ? En d'autres termes,

est-ce que les associations les plus spécifiques permettent de révéler l'effet de sens ou la valeur modale la plus fréquemment véhiculée par les modaux ?

Nous pensons que les profils combinatoires de *pouvoir* et *devoir* seront plus diversifiés que ceux de *vouloir* car le NS de ce verbe implique que les sujets soient majoritairement des êtres *animés*. Cette intuition pourrait être contredite si nos résultats démontrent que ce sont davantage les usages grammaticalisés qui apparaissent dans nos corpus.

Nous avons choisi de faire un focus particulier sur les noms en position sujet car « le syntagme nominal est souvent présenté dans les grammaires comme le prototype du sujet, soit sous une forme simple, avec un article défini [...], soit sous forme de nom propre [...], soit sous une forme plus étoffée [...] » (Blanche-Benveniste 2003 : 73-74).

Pour analyser au mieux ces sujets grammaticaux, nous regroupons les noms en position sujet dans des catégories sémantiques. En effet, il est plus pertinent d'analyser qu'un verbe s'associe davantage avec des noms d'*événements* plutôt qu'avec des *êtres animés* avant de nous attarder sur une catégorie en particulier. En ceci, nous pouvons ici bien voir que nous sommes tant dans une approche corpus *based* (nous aimerions vérifier ou infirmer les remarques des autrices qui nous précèdent concernant l'influence du sujet sur l'interprétation sémantique du modal) que *driven*, car nous allons nous focaliser surtout sur les associations les plus significatives de ces modaux avec leurs sujets afin d'en étudier les caractéristiques sémantiques.

Parmi les nombreuses classifications nominales existantes, nous avons opté pour la typologie des noms de Salvadori et al. (2021) que nous présentons dans la section suivante.

7.1.1. La typologie des noms de Salvadori et al. (2021)

Les noms ont fait l'objet de nombreuses classifications sémantiques, et ceci est sûrement dû à leur aspect « concret » des représentations mentales que nous en faisons :

Le lexique nominal a souvent été caractérisé, par opposition au lexique adjectival ou verbal, par le fait que les N pouvaient être décrits au moyen de traits "inhérents" (/animé/, /humain/, /concret/...), tandis que les verbes et les adjectifs auraient été pourvus de traits "contextuels" (/sujet animé/), /objet humain/...). (Flaux & Van de Velde 2000 : 113)

Il existe de très nombreuses et récentes classifications sémantiques des noms. Flaux et Van de Velde (2000) fournissent une typologie fondée sur une approche syntaxico-morphologique plutôt que de sémantique pure. Elles délimitent ainsi deux grandes classes qui sont des « noms véritables » *versus* « des noms dérivés à partir de verbes ou d'adjectifs ». Nous pouvons aussi

citer le projet Nomage (Balvet et al. 2011) qui se consacre aux noms déverbaux¹⁴³. Fasciolo et Lammert (2014) ont publié un numéro de *Langue française* dont les articles examinent des critères liés à certaines classes spécifiques. Une année plus tard, Huyghe (2015) coordonne lui aussi un numéro dans cette revue avec comme angle d’attaque les critères sémantiques pour déterminer les classes des noms.

Les recherches concernant certaines catégories nominales sémantiques particulières sont également foisonnantes¹⁴⁴. Le défi rencontré avec ces classifications réside dans le fait qu’elles proposent toutes des catégories, mais aucune ne dispose véritablement d’un dictionnaire ou d’une liste exhaustive regroupant les noms selon les groupes sémantiques établis. C’est pourquoi nous avons adopté une typologie qui facilite la classification de nos résultats, celle présentée dans Salvadori et al. (2021), largement inspirée de celle de Balvet et al. (2011), et donc axée sur la classification des noms déverbaux¹⁴⁵. Bien qu’elle ne fournisse pas une liste exhaustive des noms classés selon les groupes sémantiques définis, elle propose des indications sur la manière de les catégoriser au sein des ensembles établis, ce qui la distingue des autres classifications mentionnées dans cette section.

La marche à suivre pour classer les noms est détaillée dans Haas et al. (2022). Elle s’appuie sur des caractéristiques distributionnalistes et observe si tel ou tel nom est compatible « avec des constructions propres à certaines catégories sémantiques » (Haas et al. 2022 : 2). Ces tests se font à l’aide de dictionnaires et de corpus afin d’observer quels sont les emplois effectifs des entités nominales à classer.

Les classes sont au nombre de 14 et sont nommées en vert¹⁴⁶ dans l’arbre de décision¹⁴⁷ ci-dessous :

¹⁴³ Ce projet ne semble plus disponible à l’heure où nous écrivons ces lignes.

¹⁴⁴ Pour les noms abstraits : Anokhina-Abeberry (2000) ; pour les noms de sentiments et d’attitude : Anscombe (1995) ; pour les noms d’agents : El Cherif (2011) ; pour les noms d’idéauté : Flaux et Stosic (2014) ; pour les objets et les événements : Godard et Jayez (1996) et Huyghe (2012) ; pour la polysémie des noms d’affect : Goossens (2009) ; pour les noms d’activité : Haas et Huyghe (2010) ; pour les noms d’événements : Huyghe (2011, 2012, 2013) ; pour les noms d’action : Huyghe (2014) ; pour les noms d’agents et d’instruments : Huyghe et Tribout (2015) et pour les noms dénotant des humains, voir le numéro de *Linx* proposé par Schnedecker (2018). Notons que les chercheuses travaillent parfois sur des catégories similaires mais qu’elles dénomment différemment (par exemple *agent* et *humain* peuvent désigner des noms similaires).

¹⁴⁵ Leur typologie est toutefois adaptable à tout type de nom.

¹⁴⁶ Concernant les catégories figurant dans des encadrés en jaune, celles-ci représentent des possibilités de regroupements plus complexes, la granularité fine de cette classification. Nous ne prenons en compte que la granularité moyenne dans notre étude et considérons ainsi les *artefact*, *objets naturels* et *événements* comme des catégories en tant que telles. Ce choix semble se justifier car « les différences entre les annotations lexicales et contextuelles sont minimisées lorsqu’on considère des types sémantiques à granularité moyenne, dans lesquels les types complexes sont fusionnés » (Varvara et al. sous presse : xx). Pour notre type de recherche, les résultats seront ainsi plus probants avec ce grain d’analyse.

¹⁴⁷ Le T correspond à *test* dans Haas et al. (2022).

C'est seulement si le nom peut s'insérer dans les trois énoncés du test (à la place du *N*) qu'il peut être considéré comme un *animé*. Parfois, le nom doit pouvoir intégrer seulement un des énoncés du test pour être classé dans une certaine catégorie :

- T3¹⁰**
a. $N_{[humain]}$ a {fabriqué / déchiré} Dét N {à tel endroit / à tel moment}
OU
b. $N_{[humain]}$ a construit Dét N + Compl de matière + Adj de couleur
OU
c. $N_{[humain]}$ a confectionné Dét N + Compl d'arôme

Figure 13 : Test 3 de Haas et al. (2022 : 9) qui permet de déterminer si le nom rencontré est un artefact dans la classification de Salvadori et al. (2021).

Dans la figure 13, le nom doit pouvoir remplacer *N* dans seulement un énoncé pour être intégré dans le groupe des *artefacts*. Par conséquent, si le nom ne répond pas aux tests de la première case du schéma, nous passons à la seconde, et ainsi de suite, jusqu'à ce que le nom corresponde aux énoncés des tests¹⁴⁸. Nous avons ainsi nous-même formé une sorte de « dictionnaire »¹⁴⁹ qui nous permet de classer rapidement les noms figurant en position de sujet pour les verbes que nous étudions ici.

Nous avons également opté pour une granularité plus épaisse de ces catégories afin d'avoir des représentations des résultats encore plus générales, présentées dans Varvara et al. (2022)¹⁵⁰ et Varvara et al. (sous presse). Nous pouvons ainsi réduire ces 14 classes à 3 classes qui sont les noms de type *animé*, *objet* ou *événement*. Nous obtenons ainsi le tableau des classes possibles suivant¹⁵¹ :

¹⁴⁸ Concernant les termes polysémiques, nous avons décidé de les classer dans la catégorie la plus adéquate, en nous basant sur la représentation la plus fréquente de ce nom dans des échantillons tirés des corpus que nous utilisons pour cette étude.

¹⁴⁹ Il est consultable sous forme de tableur Excel dans le lien suivant et comporte près de 2696 entrées : <https://drive.switch.ch/index.php/s/anKNMbZ9NTDheNW>

¹⁵⁰ Nous n'avons malheureusement pas pu être présente pour la présentation que nous mentionnons via cette source mais nous remercions chaleureusement Justine Salvadori d'avoir partagé avec nous la classification présentée lors du colloque *20th International Morphology Meeting (IMM20)* qui s'est tenu à Budapest en 2022.

¹⁵¹ Nous présentons les catégories de la granularité moyenne selon leur ordre d'apparition dans le guide d'annotation de Salvadori et al. (2021).

Tableau 1 : Adaptation de la typologie des noms de Salvadori et al. (2021) à trois catégories de Varvara et al. (sous presse).

Granularité moyenne	Granularité épaisse
animé	animé
artefact	objet
objet naturel	objet
évènement	évènement
domaine	évènement
propriété	évènement
état	évènement
institution	animé
phénomène	évènement
objet cognitif	objet
objet financier	objet
temps	évènement
quantité	objet
maladie	évènement

Il peut être étonnant de voir figurer dans la colonne de la granularité épaisse les titres d'*animé* et *évènement* qui sont des catégories de la granularité moyenne. Il semblerait que ces noms de groupe aient été choisis car ces deux catégories de la granularité moyenne représentent le mieux ce regroupement dans la granularité épaisse. Les noms *animés* et *évènement* sont aussi plus fréquents que les autres catégories qui se trouvent regroupées sous ces étiquettes dans la granularité épaisse. Il y a ainsi des noms d'*animés* et d'*évènement* stricts (dans la granularité moyenne) et d'autres englobants (dans la granularité épaisse).

7.1.2. Préambule à la recherche

Avant d'exposer les résultats, nous tenons à faire un point quant à la présentation de ceux-ci afin de clarifier notre démarche mais aussi la lecture de cette section.

A partir des cooccurents spécifiques nominaux des verbes modaux dans un empan de 4 mots avant le pivot, nous allons passer par trois étapes d'analyse. La première consiste à étudier ces résultats avec la granularité épaisse de Varvara et al. (2022, sous presse) pour avoir un premier aperçu très général des profils combinatoires de ces trois verbes. Ensuite, ces résultats seront affinés à l'aide de la granularité moyenne de cette même catégorisation (Salvadori et al. 2021). Enfin, selon ces premières observations, nous creuserons les similarités et disparités des verbes modaux en nous focalisant sur les catégories qui montrent le plus de divergence ou de

ressemblance. Dans cette dernière étape, nous allons observer des extraits de notre corpus afin d'exemplifier des usages fréquents de ces verbes modaux selon la catégorie nominale qu'ils attirent. Ces exemples peuvent apporter un éclairage sur les données chiffrées qui montrent des attractions significatives de ces verbes avec certains types de sujet.

Les résultats sont ici présentés par corpus, en commençant par LM puis en enchaînant avec WIKI. Pour chacun des corpus, nous réalisons les étapes d'analyse susmentionnée pour trois formes de nos verbes, soit les *lemmes*, le *présent de l'indicatif* ainsi que le *conditionnel présent*. Cela nous permettra ainsi d'amener un éclairage nouveau sur les propriétés sémantiques des modaux et de constater si, par exemple, leur NS est plus ou moins présent selon la forme à laquelle ils sont conjugués.

Une fois ces profils combinatoires passés à la loupe, nous finirons par présenter les résultats des AFC et de l'analyse arborée afin de mettre en parallèle les résultats présentés dans ces premières sections.

Falloir, en tant que verbe impersonnel, ne peut pas être mis au même niveau d'analyse que les autres modaux que nous étudions ici. Nous lui consacrons cependant une section afin de voir de plus près quels sont les sujets nominaux spécifiques exprimés dans la complétive (7.1.5.).

7.1.3. Les résultats dans LM

Nous allons commencer par observer les résultats dans le corpus journalistique LM avec les différents paramètres susmentionnés.

7.1.3.1. La granularité épaisse pour les formes lemmatisées

Dans ce corpus et selon ce premier niveau d'analyse, *pouvoir* et *devoir* semblent présenter des profils combinatoires similaires par rapport à *vouloir* qui préfère largement les sujets de type animé :

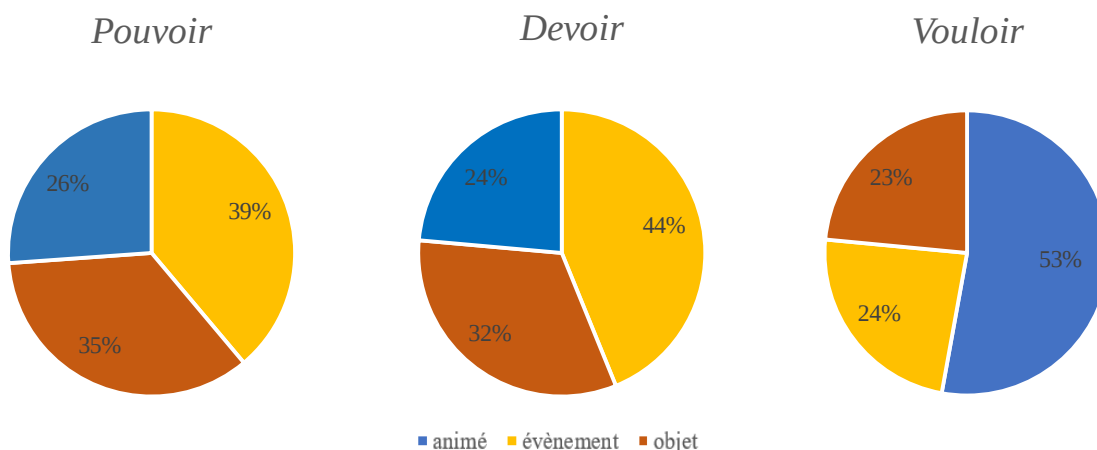


Figure 12 : Répartition des noms en position sujet selon la granularité de Varvara et al. (sous presse) dans le corpus LM.

Ces premiers graphiques démontrent une première dichotomie entre *pouvoir* et *devoir* d'un côté, et *vouloir* de l'autre. Les deux premiers ont un profil combinatoire très similaire en tant que lemme dans LM, avec une répartition plutôt égalitaire du type de nom qui se trouve en position sujet. Une prédominance pour les sujets de type *événement* est à noter de manière plus marquée du côté de *devoir*. Quant à *vouloir*, nous remarquons que ce sont bien les sujets de type *animé* qui se retrouve plus d'une fois sur deux en position sujet parmi les cooccurents significatifs de ce modal. Toutefois, cela signifie que près de la moitié des sujets significatifs ne sont pas des *animés* ce qui montre que dans beaucoup de cas, *vouloir* se grammaticalise, ou, du moins, présente des usages moins communs compte tenu de son NS. Nous allons détailler ces résultats dans les sections suivantes.

7.1.3.2. La granularité moyenne pour les formes lemmatisées

Si nous regardons désormais plus en détail les types de noms qui figurent de manière significative en position de sujet pour ces trois verbes dans LM, nous remarquons que les observations que nous avons soulevées dans la partie précédente sont confirmées :

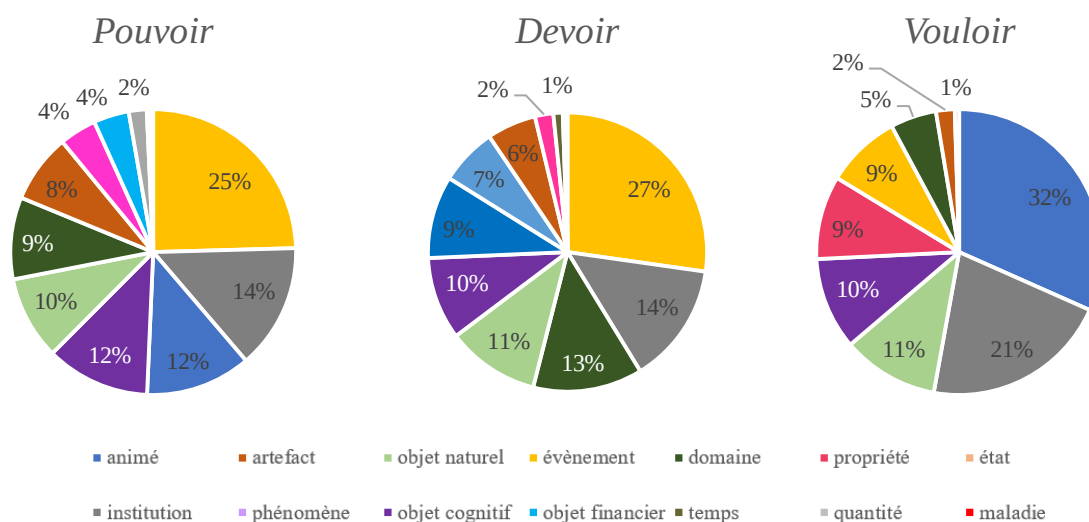


Figure 13 : Répartition des noms en position sujet pour les formes lemmatisées des modaux selon les catégories de Salvadori et al. (2021) dans le corpus LM.

Pour *vouloir*, nous comprenons que la grande importance de la catégorie des *animés* dans la granularité épaisse n'est pas seulement due aux êtres *animés* en tant que tels, mais également aux *institutions*. Nous reviendrons sur ces résultats plus en détail en 7.1.3.5.

Pour *pouvoir* et *devoir*, nous remarquons que les deux catégories qui arrivent en pole position pour ces verbes sont les mêmes, soit les noms d'*événement* et les *institutions*. Ils partagent donc également des points communs avec *vouloir* dont nous venons de relever l'importance de ces types de sujets¹⁵². Cependant, *pouvoir* et *devoir* diffèrent quelque peu sur le reste des proportions. Ainsi, *pouvoir* semble avoir plus d'attraction avec des sujets *animés* et des *objets cognitifs* alors que *devoir* apparait de manière plus significative avec des noms dénotant des *domaines* ou encore des *objets naturels*.

Dans les parties suivantes, nous allons observer de plus près ces divergences et ces ressemblances à l'aide de l'indice de spécificité et d'exemples tirés du corpus, ce qui nous permettra de comprendre en usage en quoi ces modaux se ressemblent ou divergent.

¹⁵² Ces résultats peuvent aussi être un biais du genre textuel. En effet, dans un corpus de presse nationale comme LM, les noms d'*institutions* sont fréquents car les articles traitent en grande partie de l'actualité politique. En consultant l'index hiérarchique du corpus (obtenu avec la fonction *lexique* dans TXM), nous pouvons observer que le nom *pays* est le deuxième nom le plus fréquent de ce corpus, et *gouvernement*, le dixième, tous deux des noms de type *institution* selon la typologie de Salvadori et al. (2021), ce qui corrobore ici notre hypothèse.

7.1.3.3. Les sujets de type évènement de pouvoir et devoir en forme de lemme

Nous avons pu noter que *pouvoir* et *devoir* avaient pour point commun d’attirer tous deux en majorité des noms de type *évènement*, noms qui, en tant que sujets grammaticaux, n’ont pas encore été étudiés pour l’influence qu’ils pouvaient potentiellement apporter à la lecture sémantique du modal. Nous allons donc nous pencher sur cette catégorie-là afin d’observer ce que cette association importante peut nous dire de leurs caractéristiques sémantiques :

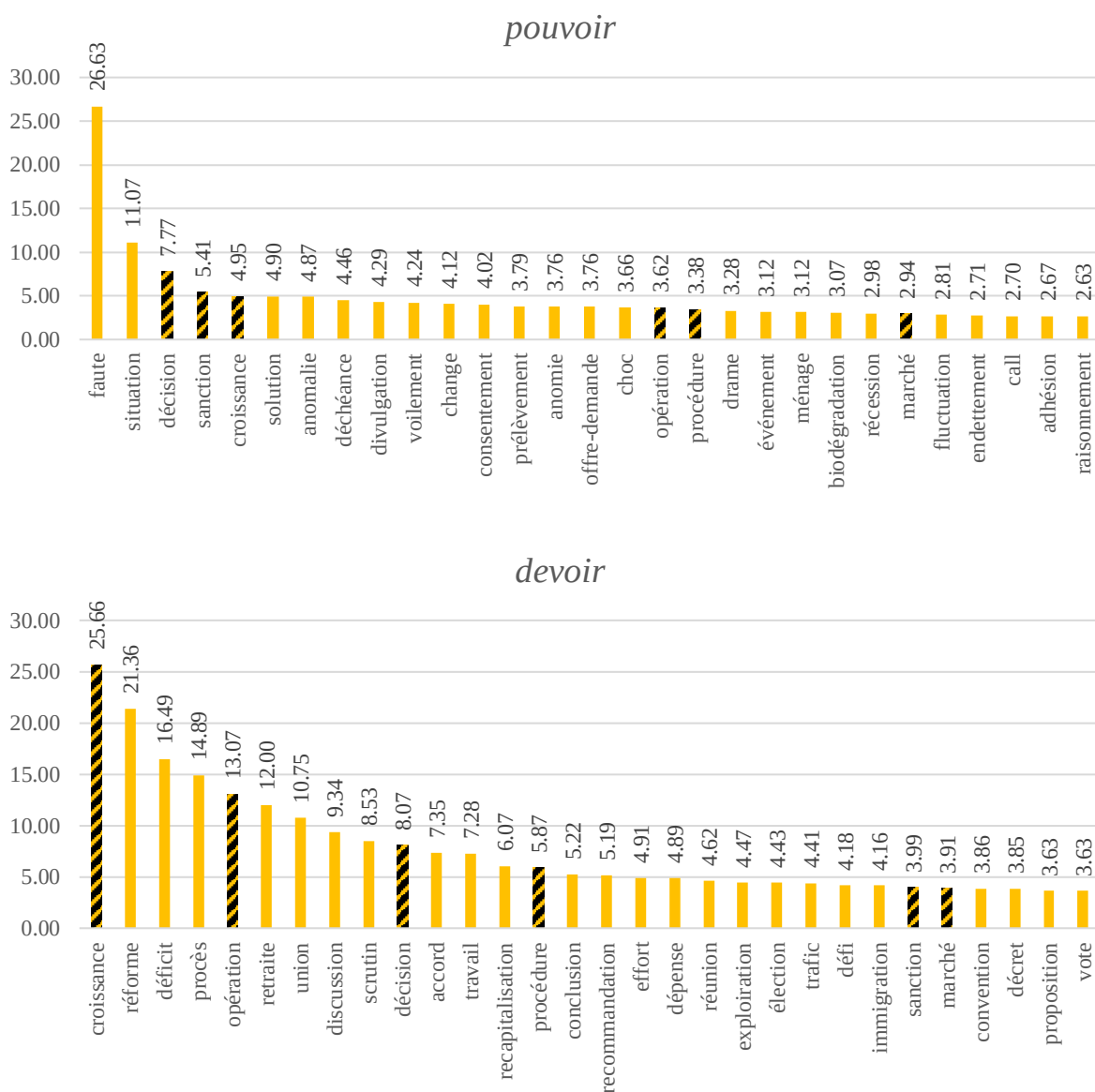


Figure 14 : Les 30 sujets nominaux de type évènement avec le plus grand indice de spécificité pour pouvoir et devoir dans le corpus LM.

Parmi les 30 sujets nominaux de type *évènement* avec l'indice de spécificité le plus haut pour ces deux verbes, nous pouvons constater qu'ils ont six sujets communs¹⁵³, soit *décision*, *sanction*, *croissance*, *opération*, *procédure* ou encore *marché*. Si l'on observe de plus près les occurrences de ces verbes avec ces sujets en particulier, nous notons que dans 40% des cas, les verbes *pouvoir* et *devoir* sont au conditionnel, dans 19% au présent, et dans les 41% restant, aux autres temps verbaux.

Les formes au conditionnel montrent pour *pouvoir* et *devoir* un usage que nous pourrions qualifier de *prédiction* :

67. Plus inquiétant : la **croissance** américaine **devrait** cette année dépasser de deux points celle de l'Europe (3, 1 % contre 1, 1 %). (LM, 05.04.2010)
68. L'**opération devrait** avoir un effet positif sur le bénéfice par action dès la première année. (LM, 15.04.2010)
69. Cette **décision devrait** permettre à Grasset de poursuivre la publication des œuvres complètes du philosophe. (LM, 16.04.2010)
70. Ces tours de vis budgétaires interviennent alors que la **croissance ne devrait** pas dépasser 1 % en 2010. (LM, 12.05.2010)
71. Du côté de Vivendi, on laisse entendre que l'**opération pourrait** être réglée dès l'été. (LM, 16.04.2010)

Cet effet de sens de *prédiction*¹⁵⁴ est attribué à *devoir* dans le discours journalistique et est défini comme « un événement¹⁵⁵ qui est/était supposé se produire » (Rossari et al. 2022a : 408). Cette définition semble pouvoir s'appliquer aux exemples 67 à 69. Cette interprétation peut également s'appliquer à des énoncés où le verbe est nié comme en 70. Enfin, nous estimons que cet effet de sens peut aussi s'appliquer au verbe *pouvoir* comme en 71. Notons également que la *prédiction* semble être liée à des éléments temporels. En effet, tant dans 67, 68, 70 que 71 figure un complément circonstanciel de temps qui permet ainsi de « projeter » l'évènement en question dans un futur plus ou moins proche.

¹⁵³ Les cooccurents communs sont représentés dans la figure 16 et les suivantes par les bâtons hachurés.

¹⁵⁴ Il nous semble intéressant de relever que dans Rossari et al. (2022a), nous mentionnions des *valeurs* de *devoir* et avons ainsi regroupé sur un même niveau la *nécessité*, la *prédiction* et la *modalité épistémique*. Ceci démontre une fois de plus l'embarras terminologique dans lequel nous nous situons lorsqu'il est question de caractériser les propriétés sémantiques des verbes modaux. La *prédiction* n'était alors pas associée aux valeurs modales épistémiques que ce verbe pouvait véhiculer mais représentait une autre valeur que ce modal pouvait avoir.

¹⁵⁵ Notons que dans Rossari et al. (2022a), la notion d'*évènement* ne faisait pas référence au sujet nominal mais au sens de l'énoncé dans sa globalité, donc en tant qu'évènement inféré comme dans : « **Les enfants devaient** défiler dans le quartier, mais les averses particulièrement fortes ont modifié les plans » (L'Est Républicain, 02.03.2010) (Rossari et al. 2022a : 408).

Cette *prédiction* semble cependant moins certaine pour *pouvoir* que pour *devoir* ce qui rejoint les analyses à leur sujet quand ils sont comparés quant à leur capacité à exprimer la valeur modale épistémique (4.6.1.). Nous pourrions ainsi nuancer cet effet de sens et parler d'*hypothèse* pour *pouvoir* comme l'illustre bien aussi l'exemple 72 :

72. « Une bombe atomique », pour les dirigeants francophones, parce que cette **procédure** **pourrait** faire imploser les fragiles équilibres de l'Etat. (LM, 27.04.2010)

L'*hypothèse* semble une lecture possible pour *pouvoir*, non seulement pour sa forme au conditionnel mais également à l'indicatif présent :

73. Des **sanctions** **peuvent** les renouer, là où la concrétisation du risque nucléaire peut au contraire les faire craquer. (LM, 17.04.2010)

Cependant, nous pensons que cette lecture d'*hypothèse* (ou dans un sens plus général de *possibilité*) n'est pas automatique pour les formes de *pouvoir* au présent avec les noms d'évènement en position sujet. Le verbe peut présenter un effet de sens de *capacité* et donc davantage se rapprocher d'une valeur modale *descriptive* :

74. Primo, les **marchés** **peuvent** s'autoréguler parce que le comportement des acteurs économiques est rationnel, puisqu'il vise à maximiser leur intérêt. (LM, 27.04.2010)

Dans 74, on interprète *pouvoir* comme attribuant une qualité (donc une capacité) aux marchés de s'autoréguler.

Pour *devoir* au présent, nous pouvons aussi noter une autre valeur modale, soit la modalité déontique :

75. Pour rénover réellement la vie publique, cette **procédure** **doit** être véritablement ouverte, commune et compétitive. (LM, 14.04.2010)

Il peut être important de relever ici le fait que *devoir* est suivi par un verbe à la forme passive¹⁵⁶, ce qui sous-entend l'action d'un agent¹⁵⁷. Ainsi, c'est peut-être l'influence d'un sujet nominal *animé* inféré qui procure à *devoir* une valeur modale déontique¹⁵⁸.

Il est ici primordial de soulever que *pouvoir* ou *devoir* n'adoptent pas automatiquement la même valeur modale ou même effet de sens avec un même sujet. Il nous semble également important de préciser que ces analyses pourront être étayées à l'aide des résultats par temps et/ou mode verbal. Nous reviendrons ainsi de manière plus détaillée sur les résultats concernant les formes au présent et au conditionnel.

7.1.3.4. Les sujets de type animé de pouvoir et de devoir pour les formes lemmatisées

Dans la figure 15, nous avons relevé qu'il y avait, certes, des ressemblances dans le profil combinatoire de *pouvoir* et *devoir* mais également des différences. Ainsi, *pouvoir* accorde une plus grande importance aux sujets de type *animés* que *devoir*. Étant donné que cette catégorie est également importante pour *vouloir* – dont les associations seront étudiées dans la section suivante – nous pensons qu'il peut être intéressant de voir en usage les rapports entre ces verbes et ce type de sujets. De plus, c'est la catégorie qui semble être le plus documentée du point de vue de l'influence qu'elle exerce sur les caractéristiques sémantiques du verbe modal dans les études précitées.

¹⁵⁶ Nous remercions ici Jérôme Jacquin qui a relevé la grande fréquence de ces tournures dans nos exemples présentés aux 11^e Journées de Linguistique de Corpus qui se sont déroulées du 4 au 6 juillet 2023 à Grenoble.

¹⁵⁷ Nous empruntons la notion de *sous-entendre* à Zribi-Hertz et Abeillé (2021) qui utilisent ce verbe pour décrire les énoncés dans lesquels l'agent n'apparaît pas mais est *sous-entendu* : « Quand [le sujet de l'actif] est absent, il est sous-entendu, comme en témoignent la présence possible d'adverbes comme *volontairement* [*Quelqu'un a volontairement cassé ce vase*] ou bien l'apparition d'ajouts infinitifs, du type *pour faire plaisir à Marie* [*La terrasse a été nettoyée pour faire plaisir à Marie*], qui impliquent bien l'existence d'un agent sous-entendu » (Zribi-Hertz & Abeillé 2021 : 211). Cette notion de *sous-entendu* peut également être mobilisée pour d'autres constructions que les énoncés passifs, par exemple, « [*dans Ces livres se vendent bien*], le vendeur n'est pas mentionné mais sous-entendu (quelqu'un vend ces livres) » (Zribi-Hertz & Abeillé 2021 : 209).

¹⁵⁸ Cette tendance pourra être vérifiée dans la section suivante qui traitera ce type de sujet pour *pouvoir* et *devoir*.

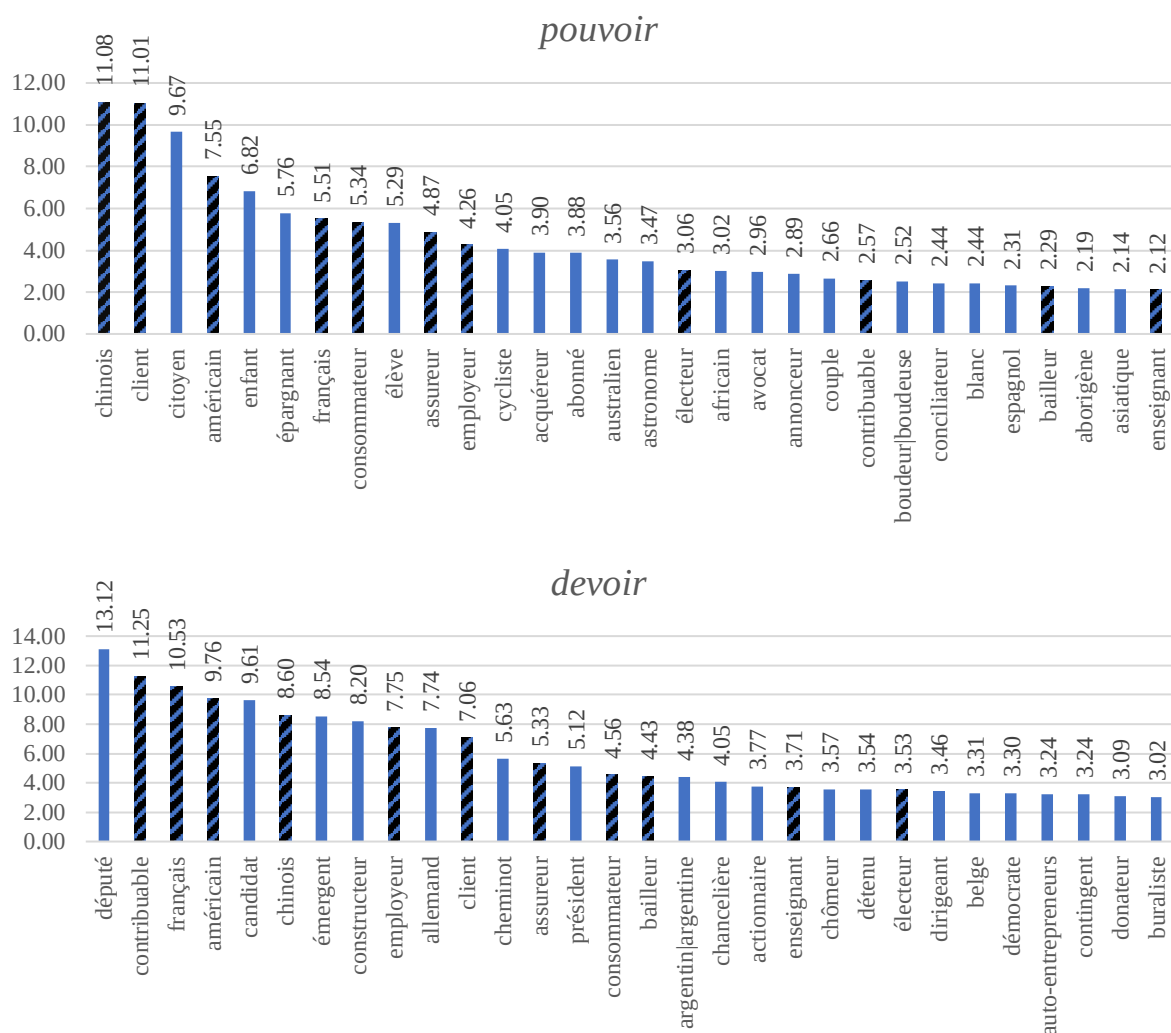


Figure 15 : Les 30 sujets nominaux de type animé avec le plus grand indice de spécificité pour pouvoir et devoir dans le corpus LM.

Même si *pouvoir* et *devoir* n'accordent pas la même importance à ce type de sujet, nous pouvons constater qu'ils attirent des noms *animés* similaires dans cette fonction grammaticale. Parmi les 30 noms en position sujet ayant le plus grand indice de spécificité, onze sont en commun chez les deux verbes : *chinois*, *client*, *américain*, *français*, *consommateur*, *assureur*, *employeur*, *électeur*, *contribuable*, *bailleur* ou encore *enseignant*, donc principalement des noms dénotant des nationalités ou alors des fonctions économiques ou politiques¹⁵⁹.

Concernant *devoir*, pour commencer, il semble que le NS de *nécessité* soit facilement récupérable avec ce type de sujet :

¹⁵⁹ A nouveau, nous pensons ici que cela est dû au corpus en tant que tel, car LM traite de ce type d'actualité de manière prépondérante comme déjà relevé dans la section précédente.

76. Au total, les **électeurs** du Sud-Soudan **doivent** donc glisser pas moins de douze bulletins dans leurs urnes respectives, alors que la région compte l'un des taux d'analphabétisme les plus élevés au monde. (LM, 14.04.2010)
77. Avocate d'Igor Soutiagine, Anna Stavitskaya a regretté que son **client ait dû** signer un document où il admet sa qualité d'« espion », ce qu'il avait toujours nié jusqu'ici. (LM, 10.07.2010)
78. S'agissant du jeu en ligne, il y a fort à parier que le **consommateur doive** encore attendre avant de pouvoir bénéficier d'une offre légale diversifiée, attractive et sécurisée. (LM, 20.04.2010)
79. Les **consommateurs doivent** prendre conscience que la mer n'est pas un libre-service [...] s'ils veulent consommer du poisson durablement ! (LM, 08.05.2010)
80. Les **contribuables doivent** estimer eux-mêmes la valeur de leurs biens immobiliers d'après leur valeur vénale. (LM, 08.05.2010)

Ces occurrences illustrent bien les remarques sur *devoir* soulevées jusqu'à présent concernant les sujets animés mais aussi les sources de l'obligation que nous avons rappelées dans l'introduction de la présente section. En effet, ce type de sujet confère à *devoir* un effet de sens de *nécessité* et/ou d'*obligation*. Notons aussi que les temps et modes verbaux sont ici variés pour cette interprétation sémantique et que c'est bien le sujet grammatical sous la forme d'un nom animé qui semble ainsi attribuer cette lecture à *devoir*.

Quant à *pouvoir*, ses interprétations sémantiques sont à nouveau plus disparates. En effet, on retrouve l'effet de sens de *possibilité* en tant que tel (81-83), mais aussi de *capacité* (84 et 85) ou encore de *permission/d'autorisation* (86 et 87) :

81. En revanche, les **contribuables pourraient** s'indigner de voir que la mobilisation des deniers publics leur a permis d'accumuler de tels gains. (LM, 06.04.2010)
82. Les **employeurs pourraient** être tentés d'embaucher des jeunes fraîchement diplômés à la place de diplômés pris au piège du chômage ou de l'inactivité depuis longtemps lorsque la reprise de l'économie s'amorcera. (LM, 15.04.2010)
83. Certains de ses **clients pourraient** aussi le trouver déplacé par les temps qui courent. (LM, 30.04.2010)
84. Dans des sondages commandés par sa société, M. Hocquemiller a ainsi pu constater que son **client pouvait** toucher les couches populaires, les émouvoir. (LM, 13.05.2010)
85. Le **client pourra** consulter sa facture avec seulement vingt-quatre heures de décalage. (LM, 11.06.2010)

86. Le **bailleur** *peut*-il demander une caution qui fasse « doublon » avec son assurance ? (LM, 16.01.2010)
87. L'**assureur** ne *peut* leur imposer ni examen médical, ni questionnaire, ni période d'attente, et ils ont droit aux mêmes taux de remboursement que les autres salariés. (LM, 16.01.2010)

Il est important de souligner que, jusqu'à présent, les remarques à propos de l'influence des sujets de type *animés* sur l'effet de sens ou la valeur modale de *pouvoir*, conférait au modal un sens de *permission*, donc une valeur modale déontique¹⁶⁰. Or, nous avons pu noter que cela ne semble pas être systématiquement le cas avec les exemples que nous venons de relever dans ce corpus. Nous remarquons ici bien avec nos exemples que ces propos sont à nuancer et que, malgré le co(n)texte, les valeurs modales et les effets de sens de *pouvoir* présentent encore des contours flous : il est ainsi peu aisé de déterminer de manière nette son interprétation sémantique, même en usage¹⁶¹.

En somme, concernant les noms *animés* en position sujet pour *pouvoir* et *devoir*, nous avons pu noter que, même si les mêmes noms pouvaient apparaître dans les cooccurents spécifiques pour l'un ou l'autre verbe, ils ne leur conféraient pas automatiquement la même valeur modale ou effet de sens.

En ce qui concerne leurs propriétés individuelles, nous avons pu apercevoir que *devoir* semble bien retranscrire son NS de la *nécessité* avec ce type de sujet. Quant à *pouvoir*, ses effets de sens et valeurs modales sont plus diversifiées. La *possibilité* est un des effets de sens envisageables dans ces co(n)textes mais cette lecture pourrait davantage être due au mode du conditionnel qu'au sujet en tant que tel. De plus, la *permission* n'est pas l'effet de sens privilégié pour cette combinaison alors que Sueur (1979) prétendait cela dans son étude présentée en 4.2.4.

7.1.3.5. Les sujets de type animé et institution avec vouloir en forme de lemme

Nous avons fait le choix de traiter *vouloir* ici à part de *pouvoir* et *devoir* car il montre un profil combinatoire différent des deux autres verbes.

¹⁶⁰ Soulignons que la valeur modale déontique s'interprète par inférence pour *pouvoir*. En effet, à partir des implicatures, nous comprenons que le *bailleur* (86) ou encore l'*assureur* (87) n'ont pas l'*autorisation* de faire telle ou telle demande.

¹⁶¹ La difficulté qu'il y a d'attribuer un effet de sens ou une valeur modale spécifique à *pouvoir* avait déjà été relevée dans notre cadre théorique (4.2.3.). Rappelons que Le Querler (2004) explique qu'« [i]l arrive souvent aussi que diverses valeurs modales se superposent » même si, selon elle, « [e]n contexte, un certain nombre d'éléments peuvent orienter l'interprétation vers telle ou telle valeur » (Le Querler 2004 : 653).

Nous avons noté dans la partie 7.1.3.2. une grande importance en position sujet, non seulement des noms dénotant des êtres *animés*, mais également des *institutions* pour *vouloir*. Cette présence importante des noms dénotant des *institutions* est due au fait que le nom *gouvernement* présente le plus grand indice de spécificité avec *vouloir* :

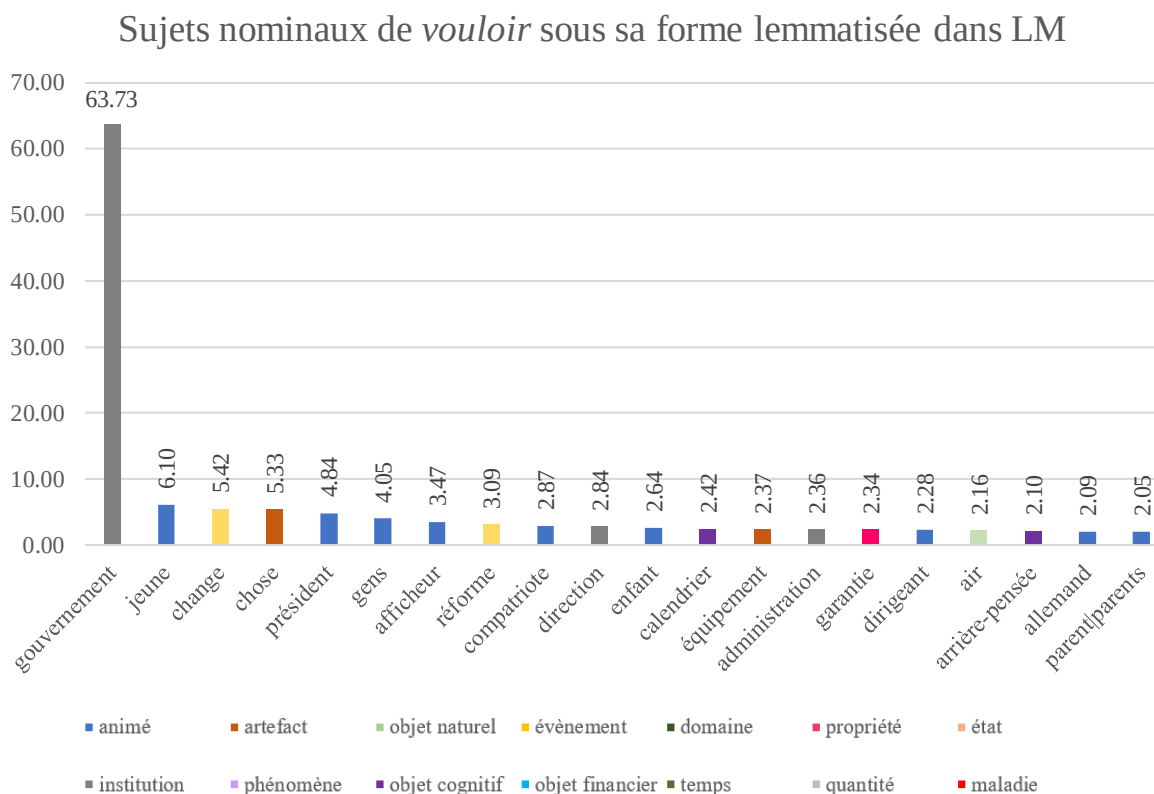


Figure 16 : Les 20 sujets nominaux avec l'indice de Lafon (1980) de *vouloir* sous sa forme lemmatisée dans le corpus dans le corpus LM.

Dans la littérature, l'association de *vouloir* avec ce type de nom n'a pas été documentée en tant que telle mais nous remarquons que ses usages sont proches de ceux avec des êtres *animés* car ils représentent des entités qui regroupent des humaines :

88. **Le gouvernement** iranien **veut** encourager quelque cinq millions de personnes à quitter progressivement Téhéran pour réduire l'impact d'un éventuel tremblement de terre. (LM, 14.04.2010)
89. **L'administration** américaine **voudrait** créer un climat propice au succès de la conférence du TNP. (LM, 14.04.2010)
90. Nous comprenons entre les lignes qu'il souhaite comme d'habitude faire payer les salariés et les retraités. On a l'impression que **le gouvernement veut** attendre l'été, peut-être la Coupe du monde de foot, pour que les gens n'aient pas le temps de réagir. (LM, 19.05.2010)

91. Dans le cas de Deepwater Horizon, la grande inconnue est l'effet des dispersants chimiques sur l'écosystème en profondeur. **Le gouvernement n'a pas voulu** se prononcer dans l'attente des conclusions de son navire océanographique, le Thomas-Jefferson. (LM, 08.06.2010)
92. Pour FO, cela passe avant tout par la suppression des objectifs personnels chiffrés et leur remplacement par des objectifs de groupe. Une mesure dont **la direction ne veut** pas entendre parler. (LM, 15.06.2010)

Dans tous ces exemples, nous retrouvons la manifestation du NS de *volonté* ou d'*intention*. Cependant, nous pouvons ici apporter quelques nuances concernant l'interprétation de l'effet de sens ou de la valeur modale véhiculée. Pour l'exemple 88, notamment, il semble que *vouloir* dénote un *ordre* de la part du *gouvernement* donc une valeur modale déontique par inférence. En effet, nos connaissances du monde mais aussi le co(n)texte nous permettent de comprendre que ce qu'un *gouvernement veut*, il l'*impose*. Cette valeur modale déontique pour *vouloir* a été attestée, notamment par Kissine (2007), mais en particulier pour les acte illocutoires directifs¹⁶². Les autres exemples ici sont aussi usités avec des *institutions* ayant du pouvoir. Cependant, nous pensons que les interprétations sont ici plus traditionnelles et cela révèle ainsi un manque de régularité dans l'effet de sens ou la valeur modale dénotée par *vouloir* avec ce type de sujet ou du moins le manque de certitude quant à la lecture sémantique de ce modal. En 89, *vouloir* exprime un effet de sens de *volonté*, atténué par l'emploi du conditionnel. En 90, nous avons également un effet de sens *volitif*, peut-être davantage lié à un *désir* puisqu'il ne semble pas s'agir d'un ordre adressé à quelqu'un. En 91, nous sommes face à une locution qui semble en voie de grammaticalisation avec le verbe *se prononcer* qui est également un verbe de communication comme *dire* avec lequel il se grammaticalise normalement. Dans cet exemple, il nous paraît cependant aisé de retrouver encore le NS de *vouloir* : en effet, ici, le gouvernement n'avait pas l'*intention* ou encore la *volonté* de révéler quoi que ce soit avant d'avoir les conclusions d'une étude. Ceci démontre que la grammaticalisation est bel et bien un processus et que *vouloir se prononcer* n'en est encore qu'aux premières phases de cette évolution vers éventuellement une locution désémantisée. En 92, nous avons la même interprétation qui émerge : *ne pas vouloir en entendre parler* tout comme *ne pas vouloir se prononcer* semblent être des cooccurrences qui représentent des premières phases de grammaticalisation. Si l'on observe désormais de plus près les occurrences avec les *animés* en position de sujet, nous pouvons constater que le NS de l'*intention* et/ou de la *volonté* est toujours récupérable :

¹⁶² Cf. la partie 4.5.3. de ce travail pour une explication plus détaillée de cette interprétation.

93. **Le président** français **veut** que Google paie une partie de ses impôts en France – c’est, par parenthèse, déjà le cas. (LM, 12.01.2010)
94. Avant de lancer des projets aux enjeux économiques majeurs - contrats, reprise d’entreprise, implantation à l’étranger... -, **les dirigeants** de sociétés **veulent** savoir ce qui les attend sur le plan juridique. (LM, 11.05.2010)
95. « Mieux vaut laisser un bibelot qu’il n’aura pas le droit de toucher », conseille Daniel Marcelli. A cet âge, très précisément, quand **l’enfant veut** se saisir d’un objet, la main en suspens, il regarde son parent pour connaître sa réaction. (LM, 19.04.2010)
96. « Pour certains [parents], il faut des mois, voire même des années pour accepter l’orientation sexuelle de leur enfant. D’autres n’y parviendront jamais », analyse M. Pinchon. **Aux jeunes** qui **veulent** le dire à leurs parents, il conseille d’abord de préparer le terrain en ayant une discussion générale sur le thème de l’homosexualité pour voir quelles réactions le sujet suscite. (LM, 17.05.2010)

Dans 93, on peut interpréter le verbe *vouloir* comme ayant une valeur modale déontique mais cette interprétation est à nouveau inférée à partir des élément co(n)textuels de l’énoncé et de prémisses. *Le président* est une figure d’autorité qui peut donner des ordres comme nous le mentionnions avec le nom *gouvernement* ci-dessus. La même lecture vaut pour 94 avec *les dirigeants*. Pour 95, avec *l’enfant*, on se retrouve plutôt avec une *demande de faire* donc une *permission* comme illustré avec l’exemple 55 dans notre partie théorique, ce qui le rapprocherait donc ici de *pouvoir*. Cette lecture est également inférée : un *enfant*, s’il *veut* faire quelque chose, doit demander une *permission* à une figure d’autorité. Concernant 96, il est intéressant de constater que nous avons le verbe *vouloir* suivi du verbe *dire* mais sans un effet de sens désémantisé ou grammaticalisé car le sujet est un *animé*. Nous retrouvons ici bel et bien une *intention de dire* avec un effet de sens de *volonté*, *désir* ou *souhait* et donc plutôt une valeur modale boulique.

En somme, il nous semble que, tant avec des sujets de type *animé*, qu’avec des sujets de type *institution*, le NS de *vouloir* est toujours décelable.

7.1.3.6. La granularité épaisse pour les formes à l’indicatif présent

Comme pour les lemmes, nous allons procéder aux trois niveaux d’analyse pour les formes de ces verbes à l’indicatif présent dans le corpus LM. La granularité épaisse montre déjà des différences notables par rapport aux profils combinatoires des lemmes présentés dans les sections précédentes :

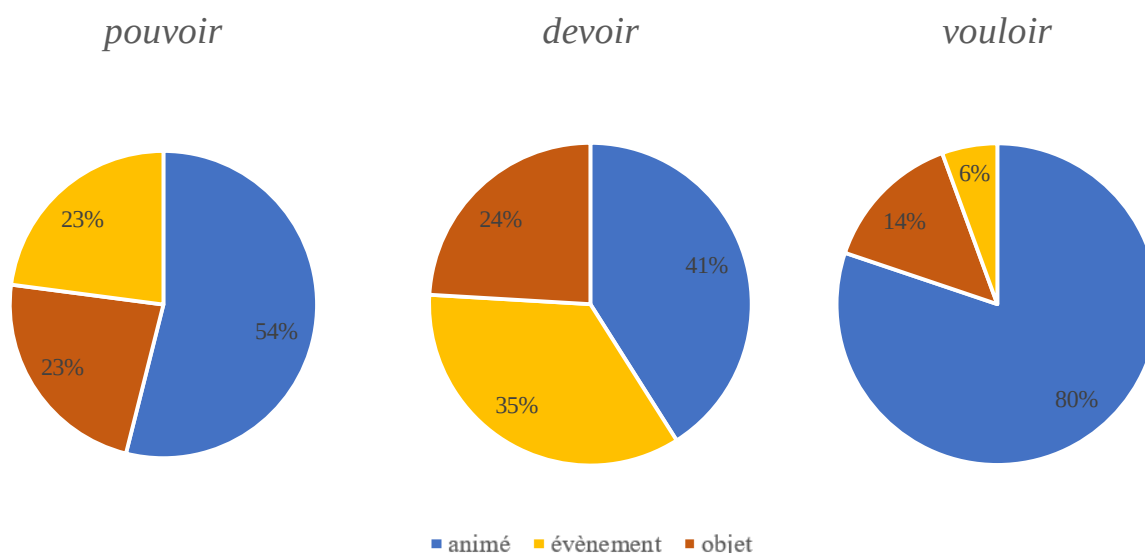


Figure 17 : Répartition des noms en position sujet pour les formes à l'indicatif présent des modaux selon la granularité de Varvara et al. (sous presse) dans le corpus LM.

Le premier constat que nous pouvons établir ici est que *pouvoir* et *devoir* ne présentent plus autant de similarités que sous leurs formes lemmatisées. Ainsi, ils attirent tous deux majoritairement des *animés* mais il semble que *pouvoir* le fasse de manière plus significative avec plus d'un sujet sur deux qui se classe dans ce groupe. Quant à *devoir*, les *animés* sont aussi les noms les plus représentés en position sujet mais de manière moins importante que pour *pouvoir*. Pour *vouloir*, la prédominance des *animés* en position sujet est plus importante encore lorsque nous prenons en compte uniquement les formes au présent dans ce corpus. Nous pouvons affiner ces remarques avec la granularité moyenne de ces catégories.

7.1.3.7. La granularité moyenne pour les formes à l'indicatif présent

Les résultats de la granularité moyenne viennent quelque peu nuancer les graphiques de la granularité épaisse. En effet, nous constatons, certes, les prédominances susmentionnées mais nous pouvons aussi noter une certaine ressemblance dans les catégories présentes dans le haut de ce classement pour ces trois verbes :

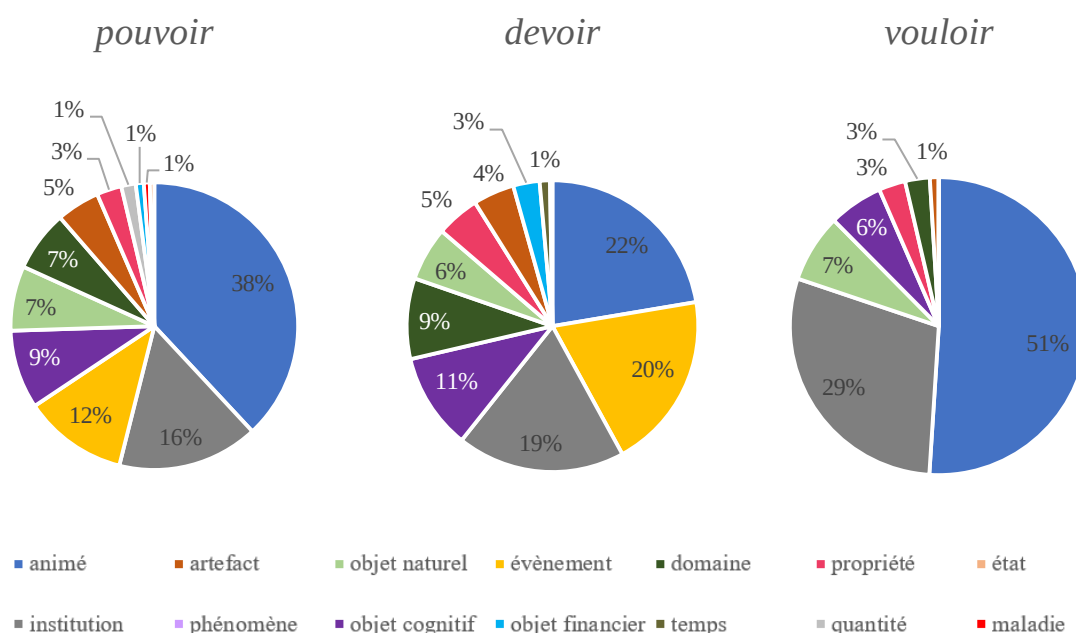


Figure 18 : Répartition des noms en position sujet pour les formes à l'indicatif présent des modaux selon les catégories de Salvadori et al. (2021) dans le corpus LM.

Comme pour les formes lemmatisées, la part des *institutions* explique l'importance de la macro-catégorie des *animés* dans la granularité épaisse. Nous pouvons aussi noter que *pouvoir* et *devoir* ne sont finalement pas si différenciés que cela car ce sont les mêmes catégories de noms qui arrivent en tête de liste pour ces deux verbes, à quelques différences près, avec, pour les deux verbes modaux, les *animés* comme étant le groupe de nom le plus représenté en position de sujet. Il n'est pas non plus surprenant de constater que ce sont ces noms-ci qui sont la part la plus importante en position sujet pour *vouloir*.

Nous allons ainsi comparer les types de sujet *animé* que ces trois verbes possèdent puis nous observerons également de plus près la catégorie des noms d'*événement* pour *devoir* car ils sont presque autant importants que les premiers en termes de proportion des résultats.

7.1.3.8. Les sujets de type animé pour les formes à l'indicatif présent : le cas de personne

Si nous observons désormais les indices de spécificité les plus importants pour les noms *animés* en position sujet, nous remarquons que *devoir* présente des points communs tant avec *pouvoir* qu'avec *vouloir* mais que ces deux derniers ne partagent que le nom *personne* comme sujet commun de cette catégorie, du moins, dans les 30 indices les plus élevés :

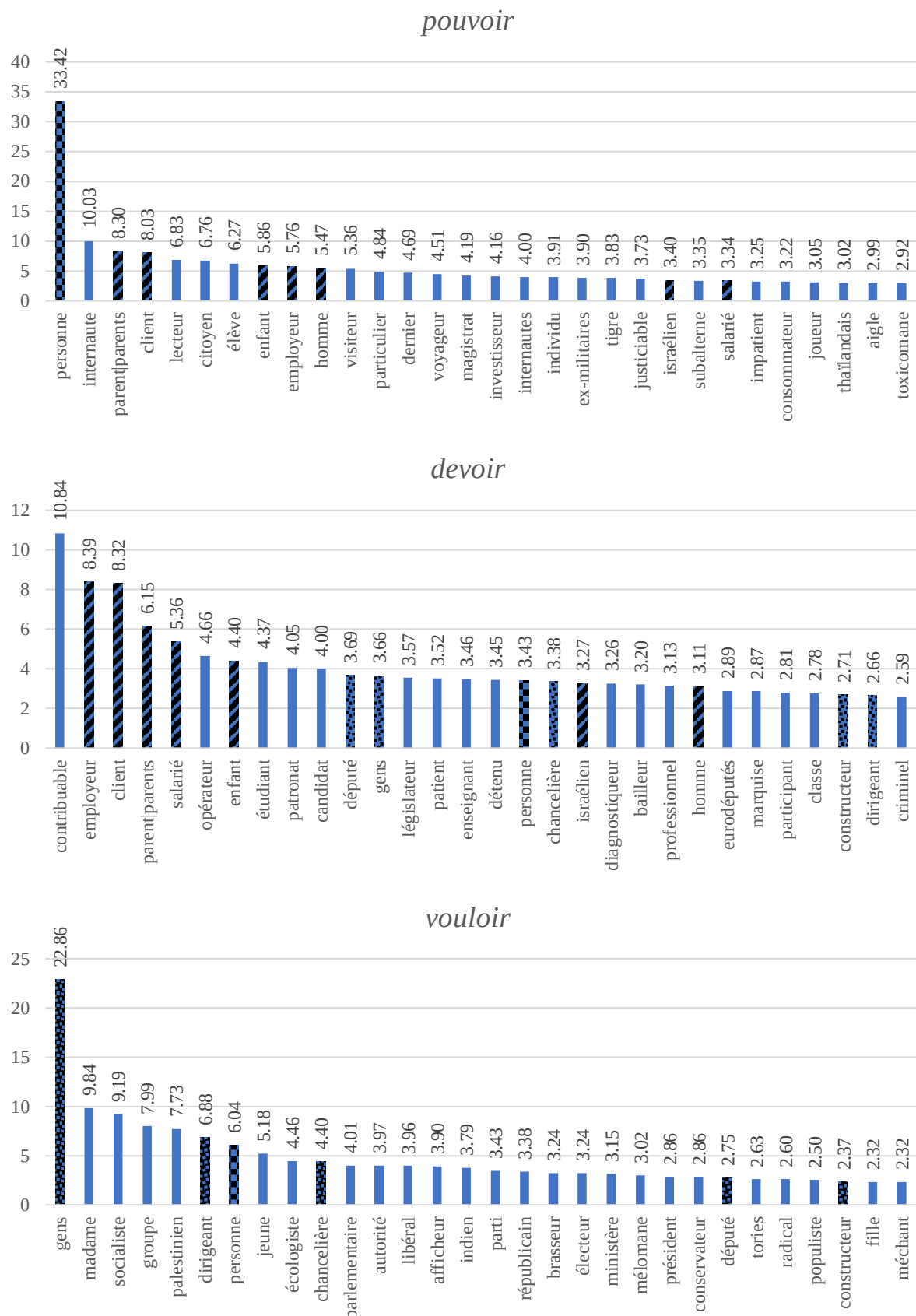


Figure 19 : Les 30 sujets nominaux de type animé avec le plus grand indice de spécificité pour pouvoir, devoir et vouloir au présent de l'indicatif dans le corpus LM.

Devoir est donc le verbe qui se distingue le moins par rapport aux deux autres verbes si on se focalise sur les sujets *animés* qui figurent en position sujet de manière significative pour lui. En effet, 13 noms sont en commun soit avec *pouvoir* soit avec *vouloir* ou encore les deux dans le cas du nom *personne*.

Si nous nous penchons sur les occurrences de *personne* avec les trois verbes au présent, nous remarquons qu'il est utilisé dans 62,5% des occurrences de *pouvoir* pour exprimer une négation, pour *vouloir* dans 60% des cas et pour *devoir* seulement 11,8% des énoncés avec ce nom sont des constructions négatives.

Pour *devoir*, que le verbe soit nié ou non, nous retrouvons aisément le NS de la *nécessité* qui peut, par exemple, exprimer une forte recommandation (97) ou alors une interdiction (98) :

97. Dans les quarante-huit heures qui suivent un rapport sexuel non protégé, un viol, une rupture de préservatif ou un échange de seringues, la **personne** exposée **doit** se rendre dans un service d'infectiologie ou aux urgences d'un hôpital pour se faire prescrire le traitement, qu'elle devra suivre pendant un mois. (LM, 29.03.2010)
98. Gratuite, cette carte est attribuée pour un an. Pour en bénéficier, une **personne** seule ne **doit** pas gagner plus de 1 040 euros par mois (1 560 euros pour un couple). (LM, 28.01.2010)

Pour *vouloir*, nous retrouvons les mêmes mécanismes en jeu, le NS est toujours encore percevable, même l'énoncé comporte une négation :

99. Un whistleblower (sonneur d'alarme) est une **personne** qui **veut** dénoncer des actes illégaux ou immoraux commis par son patron, son supérieur hiérarchique, ou par un responsable politique ou administratif, et qui possède des documents prouvant ses accusations, mais qui souhaite rester anonyme par peur des représailles. (LM, 04.12.2010)
100. Clark réussit à montrer ce que **personne** ne **veut** voir : « A l'époque, tout le monde savait qu'il y avait des incestes, de la drogue, des enfants battus, mais on faisait comme si ça n'existant pas. » (LM, 19.10.2010)

Cependant, contrairement à *devoir*, *vouloir* possède des occurrences comme en 100 où le nom *personne* occupe en fait une fonction de pronom. C'est également le cas avec le verbe *pouvoir* :

101. **Personne** ne **peut** décider de se débarrasser d'un parti politique mais **personne** ne **peut** nous empêcher d'innover (LM, 30.03.2010)
102. **Personne** ne **peut** se prévaloir de droits sur une figure historique, parce que l'histoire nous appartient à tous (LM, 19.02.2010)

Dans les deux cas, *pouvoir* semble ici véhiculer une valeur modale déontique, dans le sens d'une interdiction émanant de l'ordre de la loi voire de la morale. Cette interprétation ne paraît pas seulement être due au pronom *personne* en tant que tel mais est inférée à l'aide du co(n)texte et plus précisément avec les verbes comme *décider* et *empêcher* (101) ou encore le noms *droits* (102) qui contribuent à cette lecture de valeur modale déontique.

Quant aux énoncés dans lesquels *pouvoir* n'est pas nié, nous retrouvons à nouveau une grande diversité d'effets de sens, comme nous l'observons pour la forme lemmatisée avec les sujets *animés* plus généralement :

103. [...] le droit français de la garde à vue est conforme à la jurisprudence de la CEDH, puisque la **personne** suspecte **peut** s'entretenir avec un avocat, à la première heure. (LM, 04.03.2010)
104. Je suis inquiet de voir que des **personnes** surendettées **peuvent** se retrouver en plan de redressement personnel (PRP) pour la deuxième ou troisième fois parce que leurs dépenses dépassent structurellement le montant de leurs ressources. (LM, 22.02.2010)
105. Même les **personnes** faibles **peuvent** agir d'une façon surprenante si elles savent tirer parti d'une situation qui les pousse à se dépasser, qui les met en danger. (LM, 22.03.2010)

En 103, nous retrouvons à nouveau un usage déontique de *pouvoir* par inférence (la personne a le *droit* de s'entretenir avec un avocat), mais en 104, cela semble plutôt être une valeur modale épistémique (il y a des *chances*, des *possibilités* que cela arrive) et une valeur modale descriptive pour 105 (ces *personnes* ont la *capacité* d'agir).

De ces observations, nous pouvons ainsi conclure que *pouvoir*, à nouveau, présente le plus de variétés d'effets de sens et de valeurs modales avec un même sujet grammatical. *Vouloir*, quant à lui, ne présente pas d'effets de sens ou de valeurs modales particulières avec *personne* même lorsqu'il occupe une fonction de pronom. *Devoir* actualise, une fois de plus, fortement à son NS de nécessité dans les occurrences que nous avons observées ici.

7.1.3.9. Les sujets de type événement de devoir à l'indicatif présent

Étant donné que *devoir* semble présenter une attraction importante avec les sujets *événement* lorsqu'il est au présent et que cette attraction le différencie de *pouvoir* et *vouloir*, nous voulons ici nous focaliser sur ses usages avec cette catégorie de noms lorsqu'elle précède le modal :

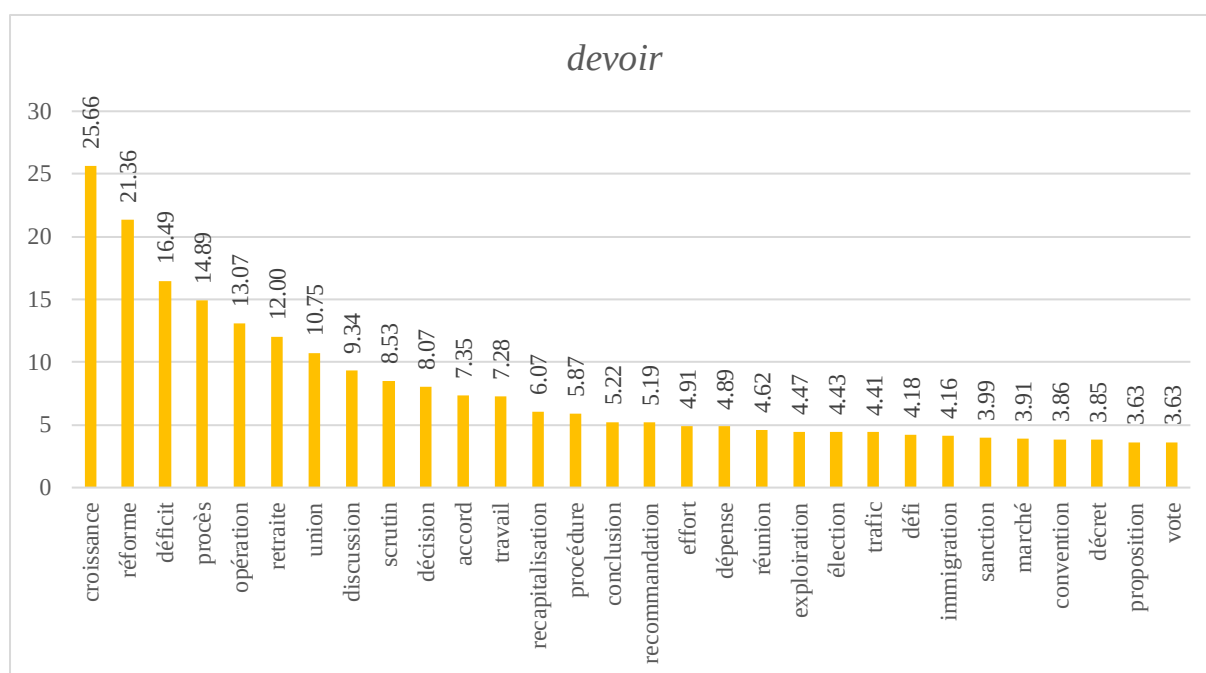


Figure 20 : Les 30 sujets nominaux de type événement avec le plus grand indice de spécificité pour *devoir* au présent de l'indicatif dans le corpus LM

Cette figure montre que ce sont des noms du domaine économique, politique ou encore juridique qui sont statistiquement associés à *devoir* dans le corpus LM. Les occurrences confirment ce que nous avons observé avec la forme lemmatisée, c'est-à-dire que la valeur modale déontique peut aisément être retrouvée lorsque *devoir* précède une tournure passive :

106. [...] on devra bien admettre que les **dépenses** fiscales **doivent** être évaluées, contrôlées, limitées dans les mêmes conditions que les crédits budgétaires eux-mêmes. (LM, 07.04.2010)
107. M. Obama a aussi essayé de dissiper l'impression que BP est aux commandes. « Chaque **décision** majeure **doit** avoir été approuvée à l'avance » par le gouvernement, a-t-il expliqué. (LM, 29.05.2010)

Ainsi, l'interprétation déontique pourrait être due au sujet *animé* même si celui-ci est seulement sous-entendu dans la forme passive. Au vu des thématiques traitées dans ces exemples, nous pouvons ici comprendre que le passif est utilisé pour « présenter la situation sans présenter la cause ni le responsable » (Zribi-Hertz & Abeillé 2021 : 219). Dans la théorie pragmatique de la politesse de Brown et Levinson (1987), il est également mentionné que les tournures passives permettent de rendre l'énoncé impersonnel et ainsi de mettre une certaine distance sur l'acte dénoté ce qui permet ainsi de maintenir les faces de la locutrice ou de l'interlocutrice (Brown & Levinson 1987 : 70). Il n'est donc pas anodin que les tournures passives soient donc

mobilisées avec *devoir* qui dénote généralement une nécessité qui peut potentiellement porter préjudice à la face.

Enfin, nous notons aussi des occurrences où c'est l'effet de sens de *prédiction* qui semble émerger, notamment avec la mention d'un repère temporel dans l'énoncé :

108. Le **procès doit** se tenir à l'automne devant la 17^e chambre du tribunal correctionnel de Paris.
(LM, 14.05.2010)

En résumé, il apparaît que l'usage du verbe *devoir* à l'indicatif présent, précédé par un sujet événement, revêt dans la majorité des cas une signification modale déontique. A ce stade de l'analyse, il devient apparent que l'influence sur l'interprétation du sens du verbe n'est peut-être pas tant attribuable au sujet nominal qu'au temps verbal utilisé, notamment, le présent de l'indicatif qui semble fréquemment conférer à *devoir* une valeur modale déontique.

7.1.3.10. La granularité épaisse pour les formes au conditionnel présent

Nous allons désormais passer aux résultats du conditionnel présent. La granularité épaisse de Varvara et al. (sous presse) révèle déjà des profils combinatoires différents de ceux que nous avons pu observer avec les formes lemmatisées ou au présent de l'indicatif :

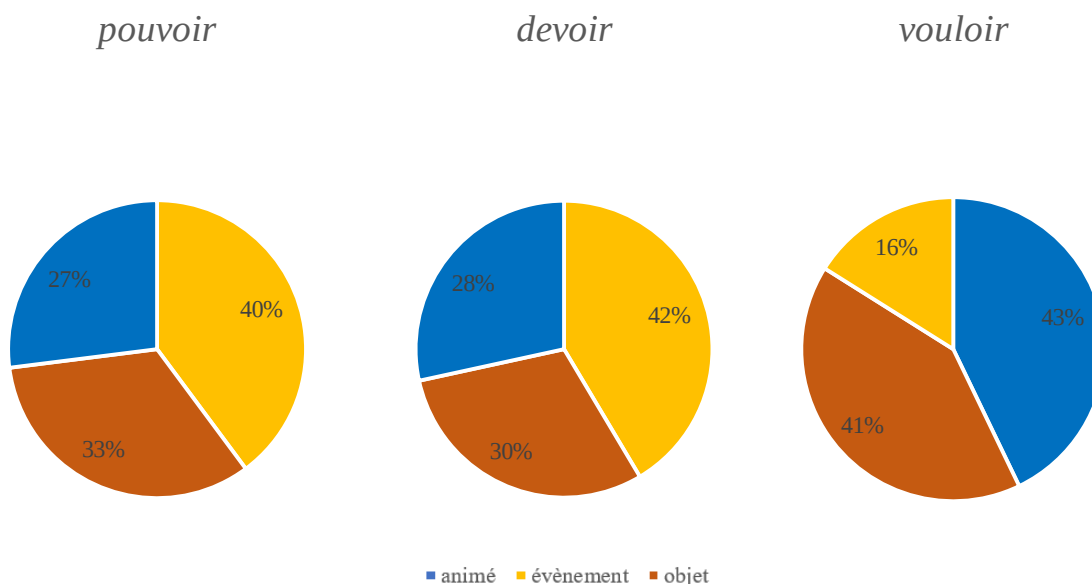


Figure 21 : Répartition des noms en position sujet pour les formes au conditionnel présent des modaux selon la granularité de Varvara et al. (sous presse) dans le corpus LM.

Les profils combinatoires de *pouvoir* et *devoir* se ressemblent à nouveau beaucoup, mais contrairement aux formes lemmatisées, la catégorie des *animés* est cette fois-ci la moins importante. Comme pour les lemmes, ce sont les noms de type *événement* qui sont le plus

représentés parmi les cooccurents significatifs de ces verbes en position sujet. *Vouloir*, quant à lui, a toujours une préférence pour les sujets nominaux *animés* mais ils sont suivis de très près par les noms classés comme *objet*. Le verbe pourrait ainsi avoir tendance à se grammaticaliser plus facilement à ce mode puisqu'il n'attire plus de manière prépondérante des sujets propres à son NS.

Ce premier aperçu des profils combinatoires de ces verbes au conditionnel présent dans LM montre des particularités qu'il n'aurait été possible de relever avec les formes lemmatisées seules. Nous allons désormais voir de plus près la composition de ces catégories.

7.1.3.11. La granularité moyenne pour les formes au conditionnel présent

La granularité moyenne nous permet, une fois de plus, d'affiner les résultats de notre première vue d'ensemble sur les profils combinatoires des verbes modaux étudiés ici.

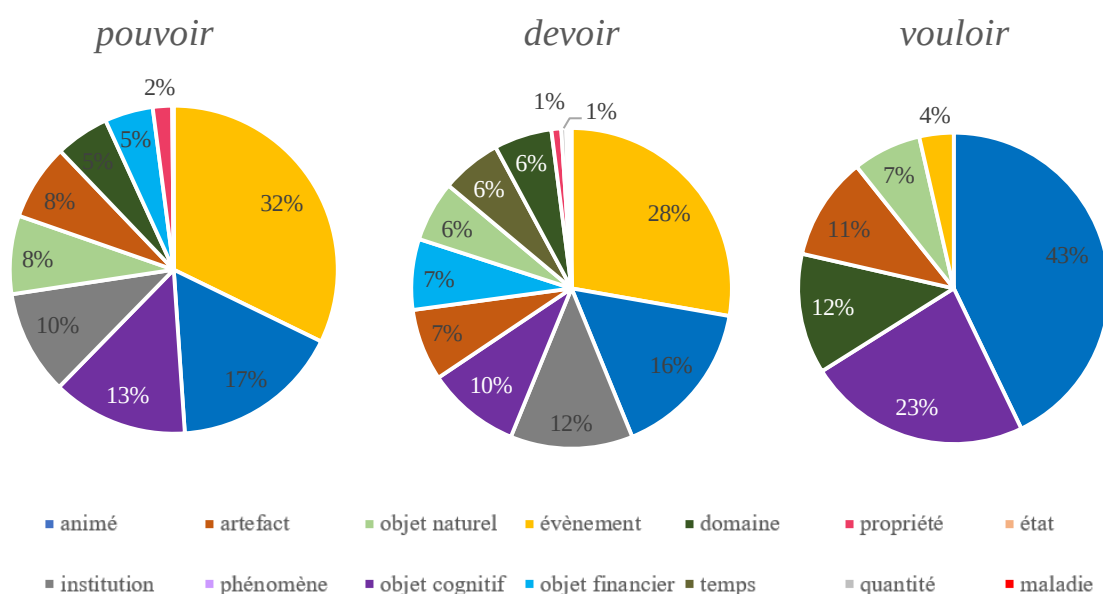


Figure 22 : Répartition des noms en position sujet pour les formes au conditionnel présent des modaux selon les catégories de Salvadori et al. (2021) dans le corpus LM.

Comme pour les graphiques précédents, les quatre catégories les plus représentées pour *pouvoir* et *devoir* sont identiques mais elles changent quelque peu d'ordre d'importance selon le verbe. Concernant *vouloir* les résultats sont relativement plus surprenants ici au vu des connaissances que nous avons pour l'heure à son propos. En effet, même si les noms *animés* représentent encore la part la plus importante des noms en position sujet significativement associés à ce verbe, presque un quart des résultats représentent des *objets cognitifs* dans ce corpus lorsque le verbe est au conditionnel.

Au vu de ces résultats, nous allons observer de plus près cette catégorie avec *vouloir* et également nous pencher sur les noms de type *événement* pour *pouvoir* et *devoir*.

7.1.3.12. Les sujets de type objet cognitif de vouloir au conditionnel présent

Alors que jusqu'à présent la prédominance pour les noms *animés* en position sujet confirmait le NS d'*intention/volonté* de *vouloir* dans nos données, il semblerait qu'une autre catégorie ait son importance lorsque ce verbe est conjugué au conditionnel présent dans le corpus de presse : les *objets cognitifs*.

Les sujets statistiquement attirés par *vouloir* au conditionnel présent et étiquetés sous cette catégorie sont au nombre de quatre :

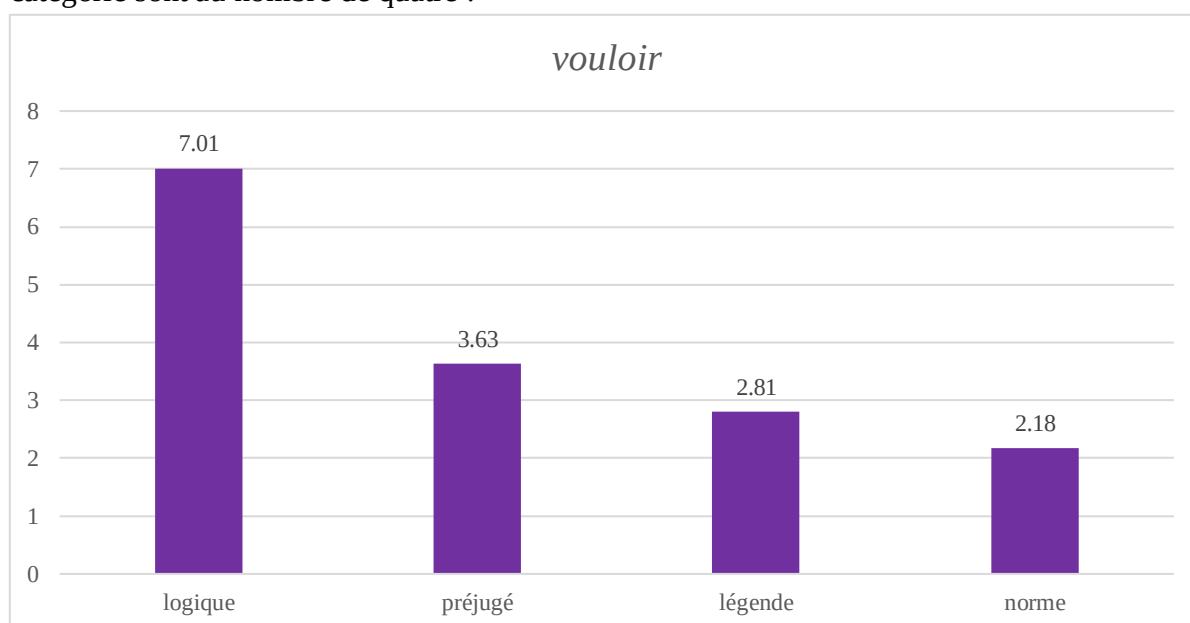


Figure 23 : Les 4 sujets nominaux de type objet cognitif pour *vouloir* au conditionnel présent dans le corpus LM selon l'indice de spécificité.

A priori, rien ne regroupe ces quatre noms hormis leur étiquette d'*objet cognitif* dans la typologie de Salvadori et al. (2021). Cependant, si nous observons de plus près les occurrences de *vouloir* au conditionnel avec ces noms, nous remarquons que pour *logique* et *norme*, le verbe permet d'exprimer des règles qui devraient être suivies ou réalisées mais qui ne le sont pas :

109. Dans le cadre du grand emprunt, de cinq à dix pôles universitaires doivent bénéficier de 7,7 milliards d'euros. « Beaucoup, mais peu par rapport à l'ampleur des problèmes », a estimé Jean-François Méla, ancien président de Paris-XIII-Villetaneuse, dans un chat, mercredi 13 janvier. Pour lui, « la **logique** » **voudrait** « que ces pôles soient relativement réduits autour des plus grandes universités actuelles », « celles qui concentrent l'essentiel des grands chercheurs des grands organismes ». Mais ce serait, à ses yeux, « une erreur magistrale,

d'un point de vue économique, et encore plus d'un point de vue démocratique, que de réduire nos préoccupations au développement de quelques pôles de recherche dans l'espoir d'avoir davantage de brevets et d'innovations en France ». (LM, 15.01.2010)

110. Une norme exigeant d'émettre un niveau minimum de sons pour avertir les aveugles est à l'étude. Pour l'instant, un seul pays au monde, le Japon, a mis en place une réglementation sur les émissions sonores des véhicules. Dans ce contexte, les constructeurs préfèrent anticiper et ont commencé à réfléchir activement à la façon d'ajouter du bruit lors du déplacement du véhicule. Mais lequel ? « C'est compliqué. La **logique voudrait** en effet que l'on rende du son aux véhicules qui n'en ont pas assez, explique Vincent Roussarie, responsable acoustique de l'entité Perceptions et facteurs humains chez PSA Peugeot Citroën. Mais nous sommes devant une double contrainte : si notre but est de pallier le risque dû au silence, il n'est pas question de créer une quelconque gêne urbaine. » (LM, 29.09.2010)
111. Marie-Pascale Robert-Rolin, une des généralistes en colère du parc naturel, qui travaille sur des secteurs de garde fusionnés, témoigne de l'étirement de son territoire : elle doit parfois parcourir des distances de 45 kilomètres sur des routes étroites et verglacées une bonne partie de l'hiver. Il est « bien difficile de respecter la **norme** qui **voudrait** que les secours soient à 20 minutes des malades. Agrandir les secteurs de garde et enlever des indemnités d'astreinte ne va pas encourager les jeunes à s'installer, comme le prétend l'administration, bien au contraire ». (LM, 06.01.2010)

Concernant les noms *légende* et *préjugé*, ils relèvent d'un *a priori* qui lui non plus ne se retrouve pas réalisé avec *vouloir* au conditionnel :

112. Pour donner corps à une **légende** qui **voudrait** qu'un vieil homme ait payé un marin pour coucher avec sa femme et lui donner un héritier, un riche marchand aigri loue les services d'une jeune femme qu'il jette dans les bras d'un matelot de passage. Mais la nuit ne se déroule pas comme il l'avait espéré... (LM, 22.11.2010)
113. L'enquête du CREN révèle un paradoxe : les mauvais élèves, de même que les filles en général, trichent moins que les autres. Et c'est l'un des enseignements et des vertus de cette étude que de malmener un **préjugé** tenace qui **voudrait** que seuls les élèves en difficulté aient recours à cet usage que réprouve la morale éducative. (LM, 27.12.2010)

A notre connaissance, cet usage de *vouloir* n'a pas été relaté jusqu'à présent alors qu'il semble fréquent. Dans les exemples 109 à 113, nous pouvons ici considérer *vouloir* comme un modalisateur dans sa définition la plus stricte, c'est-à-dire un comme un marqueur permettant

d'exprimer une « distance énonciative » (Kerbrat-Orecchioni 1980 : 32). En effet, la locutrice prend de la distance par rapport à son *dire* en utilisant le verbe *vouloir* au conditionnel. Cette distanciation est également possible car, tout comme pour les tournures passives, le sujet actif de la volonté n'est finalement pas mentionné. Il est en quelque sorte sous-entendu dans un emploi métonymique de ce type de sujet. Ces cotextes semblent favorables à une distanciation, étant donné que l'évocation des normes ou des préjugés de manière assertive dans une presse nationale – qui se doit d'être neutre – pourrait potentiellement être perçue de manière défavorable. Ce procédé pourrait donc être également rapproché d'un ménagement de la face au sens de Brown et Levinson (1987). Mentionnons donc que cet usage de *vouloir* est fréquent dans ce corpus de presse alors qu'il n'a pas été relaté dans les études à son sujet.

7.1.3.13. Les sujets de type événement de pouvoir et devoir au conditionnel présent

Les sujets de type *événement* semblent à nouveau les plus importants pour *pouvoir* et *devoir* dans ce corpus. En observant de plus près les noms qui composent cette catégorie et qui apparaissent de manière statistiquement significative en position de sujet pour ces deux verbes, nous pouvons remarquer que ce sont à nouveau des noms du domaine politico-économique qui ressortent parmi ceux ayant un indice de spécificité élevé :

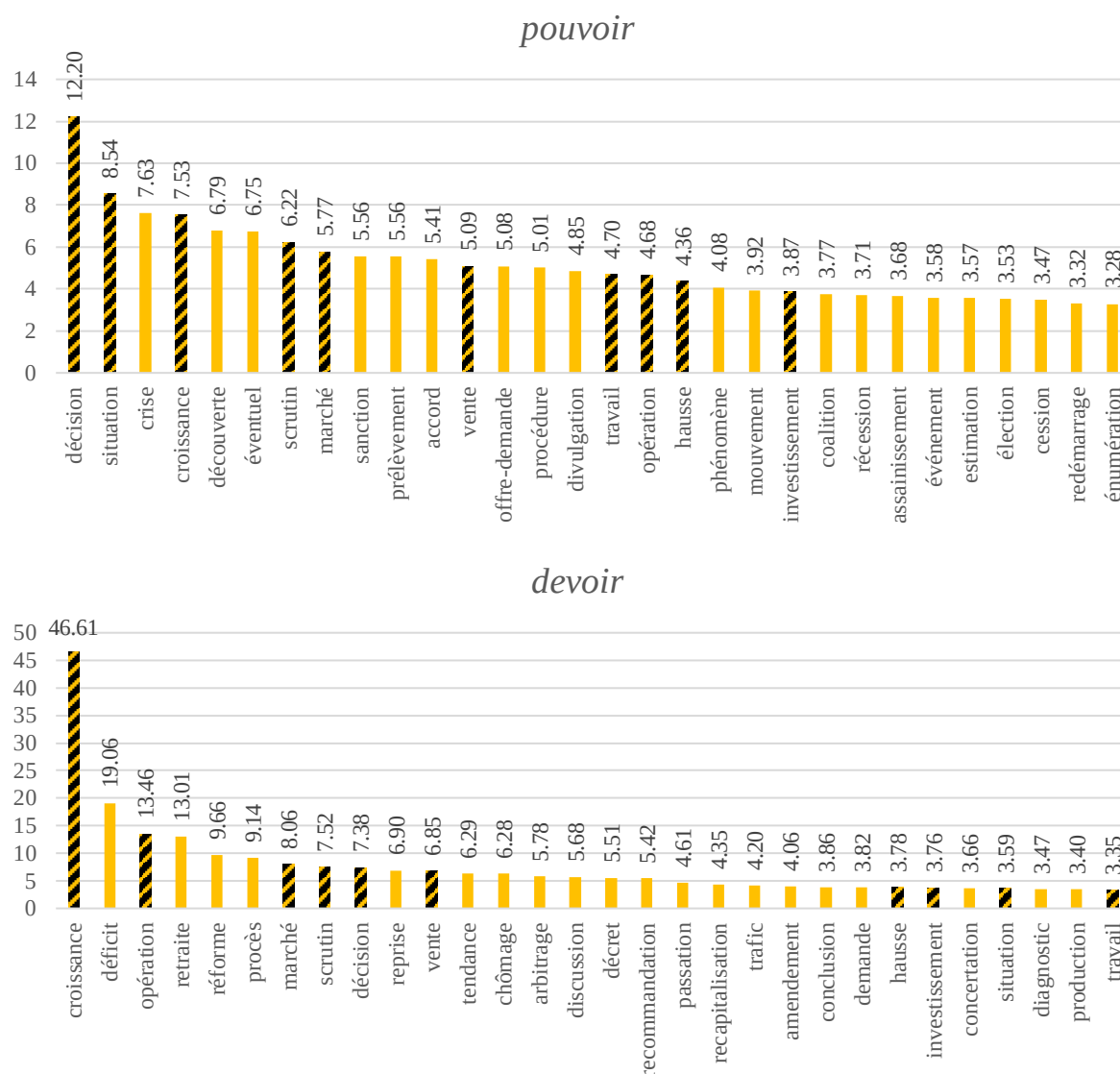


Figure 24 : Les 30 sujets nominaux de type événement avec le plus grand indice de spécificité pour pouvoir et devoir au conditionnel présent dans le corpus LM.

Ceci est encore plus marquant avec la variante du conditionnel que ça ne l'était pour les formes lemmatisées. On y trouvait aussi des termes plus génériques comme *faute*, *situation* pour *pouvoir* ; *devoir* présentait déjà cette tendance à s'associer au lexique du politique et de l'économique comme nous l'avons soulevé avec figure 16.

Parmi ces noms en position sujet, dix sont en communs pour les deux verbes, ce qui représente un chiffre plus élevé que pour les formes lemmatisées. Il s'agit ici de *décision*, *situation*, *croissance*, *scrutin*, *marché*, *vente*, *travail*, *opération*, *hausse* et *investissement*.

En observant de plus près les cotextes d'utilisation de ces noms avec *pouvoir* et *devoir* au conditionnel, nous remarquons qu'il s'agit, comme ce que nous supposions pour les formes lemmatisées, d'une expression d'une *probabilité* pour *devoir* (114-116) et d'une *hypothèse* ou d'une *probabilité moins forte* pour *pouvoir* (117-119) :

114. La **hausse** du chômage **devrait** nettement s'atténuer cette année, mais il faudra attendre 2011 pour voir la courbe s'inverser, selon des prévisions de Pôle emploi, publiées jeudi 10 juin. (LM, 11.06.2010)
115. Le cancer fait davantage de ravages dans les pays à revenu bas ou intermédiaire que dans les pays riches, et cette **situation devrait** s'aggraver d'ici 2030. (LM, 20.08.2010)
116. La fin de ces programmes d'incitations à l'achat aura des conséquences ailleurs aussi. En Allemagne, par exemple, où les **ventes devraient** atteindre, en 2009, 3, 8 millions d'unités, contre 3, 1 millions en 2008, les experts attendent une année « très rude », avec des ventes comprises entre 2, 75 millions et 3 millions. (LM, 06.01.2010)
117. Maire de Chicago depuis 22 ans, Richard M. Daley [...] a annoncé, mardi 7 septembre, à la surprise générale, qu'il ne souhaite pas se représenter aux élections de février pour un septième mandat. Cette **décision pourrait** provoquer un remaniement au sein de l'administration Obama, le secrétaire général de la Maison Blanche, Rahm Emmanuel, étant intéressé par le poste. (LM, 10.09.2010)
118. Lors du débat sur la loi, le sénateur démocrate Russ Feingold estimait que cette **opération pourrait** aussi « façonner une Africom qui travaille à la fois pour l'Afrique et pour les intérêts sécuritaires américains ». (LM, 28.07.2010)
119. La stévia, un édulcorant tiré de *Stevia rebaudiana*, plante originaire d'Amérique du Sud, a fait une entrée en force sur le marché alimentaire américain : selon le cabinet Mintel, le marché a déjà atteint 500 millions de dollars (371 millions d'euros) après le feu vert des autorités sanitaires du pays en décembre 2008 : les **ventes pourraient** atteindre 10 milliards de dollars en quelques années. (LM, 30.03.2010)

Comme susmentionné, ces effets de sens pourraient davantage être dus au mode du conditionnel qu'aux sujets en tant que tels.

Toutes les remarques que nous avons relevées pour le corpus LM sont valables uniquement pour ce corpus, *a priori*. Avec le corpus WIKI, nous allons désormais contraster les résultats que nous avons obtenus avec le comportement de ces verbes modaux dans un texte. Cette dimension comparative nous permet d'avoir un aperçu sur des potentielles régularités que les verbes modaux peuvent adopter, quel que soit le texte dans lequel ils sont usités, ou alors, de révéler des usages propres à un genre ou champ générique en particulier.

7.1.4. Les résultats dans WIKI

Comme pour le corpus précédent les trois niveaux d'analyse seront examinés tout comme les différents temps et modes verbaux sur lesquels nous nous sommes appuyée jusqu'à présent.

7.1.4.1. La granularité épaisse pour les formes lemmatisées

La granularité épaisse des catégories nominales de Varvara et al. (sous presse) nous permet de relever des différences entre les verbes mais aussi des disparités par rapport à ce que nous avons soulevé pour le corpus LM :

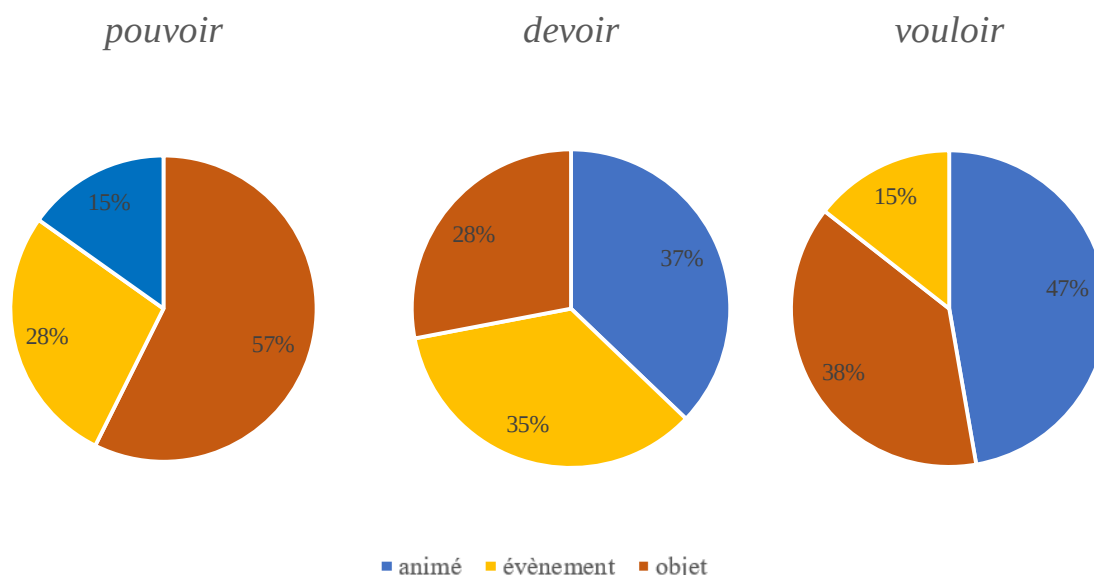


Figure 25 : Répartition des noms en position sujet pour les formes lemmatisées des modaux selon la granularité de Varvara et al. (sous presse) dans le corpus LM.

Nous remarquons que *pouvoir* et *devoir* sont ici très différents alors que leurs profils combinatoires étaient équivalents dans LM. Dans WIKI, *devoir* ne présente pas une préférence marquée pour une des trois catégories, ce qui était également le cas dans LM. Pour *pouvoir*, la donne est ici différente : plus d'un sujet nominal sur deux significativement attiré par *pouvoir* figure dans la catégorie des *objets*.

La catégorie des *animés* est celle qui est la moins représentée pour *pouvoir* sous sa forme lemmatisée alors que c'est celle qui est la plus marquée pour *devoir*, suivie de près par les noms d'*événement*. Pour *vouloir*, la préférence pour les *animés* est plus prononcée, ce qui montre à nouveau que les usages grammaticalisés sont moins importants que ceux dans lesquels il exprime véritablement son NS. Cependant, cette propension se trouve légèrement moins accentuée que dans le corpus LM (figure 14). Pour ce verbe, la catégorie des noms d'*objets* représente plus d'un tiers des occurrences des noms en position sujet pour ce verbe, ce qui est aussi plus important que dans LM.

7.1.4.2. La granularité moyenne pour les formes lemmatisées

Le niveau d'analyse de granularité moyenne nous permet d'éclairer les résultats obtenus dans la section précédente.

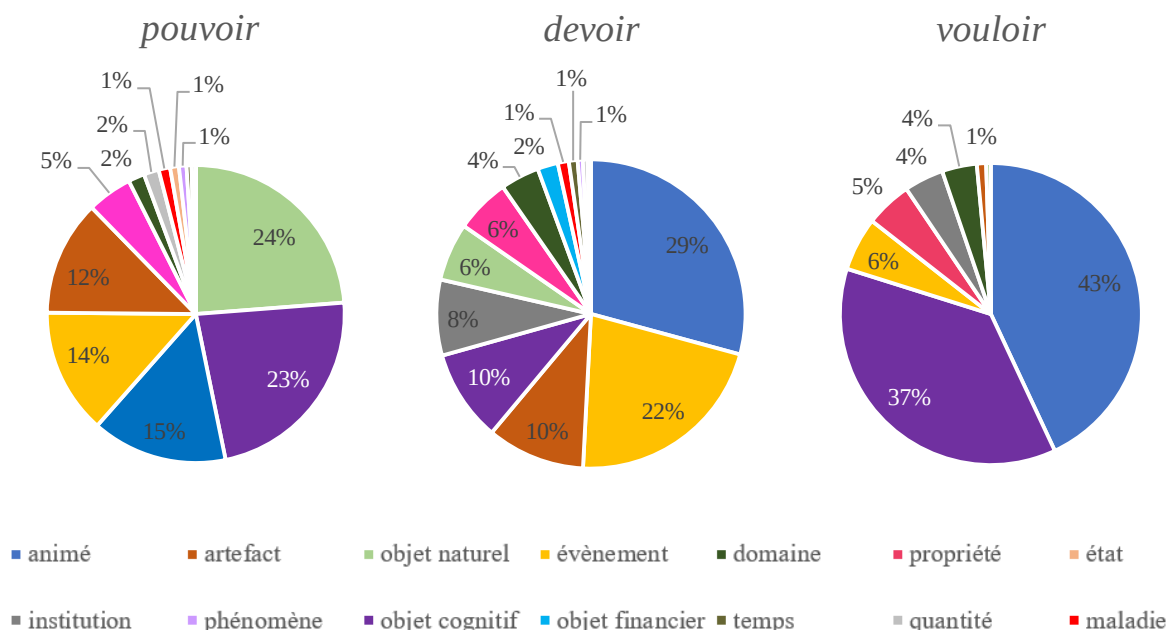


Figure 26 : Répartition des noms en position sujet pour les formes lemmatisées des modaux selon les catégories de Salvadori et al. (2021) dans le corpus WIKI.

Nous remarquons ainsi que les *objets* pour *pouvoir* sont tant des *objets naturels* que *cognitifs*. Pour *vouloir*, ce sont à nouveau les *objets cognitifs* qui représentent une part importante de ses sujets nominaux possédant un indice de spécificité significatif. Quant à *devoir*, ce sont donc les sujets *animés* qui sont les plus représentés pour sa forme lemmatisée dans WIKI.

Ce sont ces trois particularités que nous allons observer de plus près : les deux types d'*objets* mentionnés pour *pouvoir*, les *objets cognitifs* pour *vouloir* et enfin les sujets *animés* pour *devoir*.

7.1.4.3. Les sujets de type objet naturel et cognitif de pouvoir sous sa forme lemmatisée

Nous avons choisi de représenter les noms d'*objets naturels* et *cognitifs* en position sujet avec un indice de spécificité significatif pour *pouvoir* dans un seul et même graphique étant donné que leur répartition était quasi égale dans la figure précédente :

pouvoir

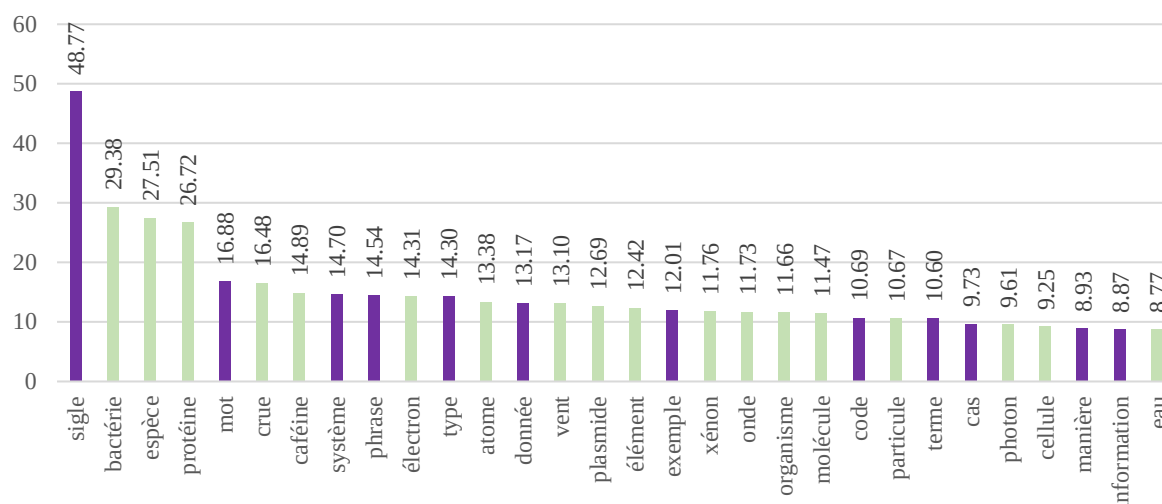


Figure 27 : Les 30 sujets nominaux de type objet naturel et cognitif avec le plus grand indice de spécificité pour pouvoir sous sa forme lemmatisée dans le corpus WIKI.

En usage, la plupart des exemples avec ces types de sujet dénotent des vérités générales et donc ce que nous pourrions qualifier comme étant l'expression d'une valeur modale descriptive :

120. Les **crues** *peuvent* être importantes compte tenu de la taille déjà grande du bassin versant. (WIKI)
121. Plusieurs organismes se proposent d'évaluer le niveau de menace sur les différentes espèces (UICN, COSEPAC, etc.) et une **espèce** *peut* donc être considérée comme « en danger critique » par l'un mais pas par les autres. (WIKI)
122. La **bactérie** symbiote *peut* fortement modifier le comportement de son hôte (et même le sexe-ratio dans certains cas). (WIKI)
123. Plusieurs ouragans frappent chaque année les côtes du golfe du Mexique et de la mer des Caraïbes, avec des **vents** violents qui *peuvent* dépasser les 200 km/h et provoquent des dégâts importants aux hôtels et habitations de la région mais aussi mettre en péril la vie des habitants. (WIKI)
124. Amaounet « La Cachée » est une déesse de la mythologie égyptienne. Elle représente le **vent** qui ne *peut* pas être vu mais qui *peut* être senti. (WIKI)
125. En effet, un **système** de composition *peut* créer différents genres musicaux. (WIKI)
126. Ces **composés** *peuvent* avoir des effets bénéfiques, lorsqu'ils sont consommés en petites quantités, ou avoir des effets néfastes lorsqu'ils sont consommés en grande quantité. (WIKI)
127. BNF est un **sigle** qui *peut* se référer à [...] (WIKI)
128. Les **phrases** autoréférentes *peuvent* être paradoxales : ainsi : « Cette phrase est un mensonge » (paradoxe d'Épiménide) ne peut être classée vraie ou fausse. (WIKI)
129. Le **mot** Veda *peut* être traduit par « connaissance ». (WIKI)

En somme, dans tous ces exemples, nous pouvons constater qu'il s'agit d'une description d'un *état de fait* pour l'objet en question et non pas d'une *possibilité* pour cet objet. *Pouvoir* semble ainsi, dans ce corpus, adopter une interprétation de valeur modale descriptive avec les *objets* en position sujet.

Ce propos est quelque peu à nuancer avec les occurrences dans lesquelles *pouvoir* est au conditionnel qui procurent à ce modal davantage un effet de sens d'*éventualité*, et donc une valeur modale épistémique :

130. Si une espèce comme le tigre à dents de sabre semble avoir naturellement disparu, **de nombreuses espèces *pourraient*** avoir brutalement disparu du fait de la chasse, bien avant l'extension de l'agriculture et des villes, dans l'hémisphère nord, en Australie et sur un certain nombre d'îles. (WIKI)
131. Les anarchistes se distinguent de la vision marxiste d'une société future en rejetant l'idée d'une dictature du prolétariat qui serait exercée après la révolution par un pouvoir temporaire : à leurs yeux, un tel **système ne *pourrait*** déboucher que sur la tyrannie. (WIKI)

Cette analyse sera à nuancer avec les résultats que nous étudierons avec le présent de l'indicatif ainsi que le conditionnel présent. Toutefois, ces remarques rejoignent les observations que nous avons pu mener jusqu'alors quant à l'interprétation sémantique de *pouvoir* : celle-ci ne semble pas dépendre du sujet grammatical – même si celui-ci peut donner quelques indices sur l'interprétation la plus probable – mais davantage du temps et/ou du mode verbal du modal.

7.1.4.4. Les sujets de type objet cognitif de vouloir sous sa forme lemmatisée

Pour *vouloir*, les noms de type *objets cognitifs* représentent une part importante des sujets dont la cooccurrence est significative. Le seul qui apparaît en commun avec ceux observés pour les

formes au conditionnel dans LM est le nom *légende* qui représente ici le nom possédant l'indice de spécificité le plus important :

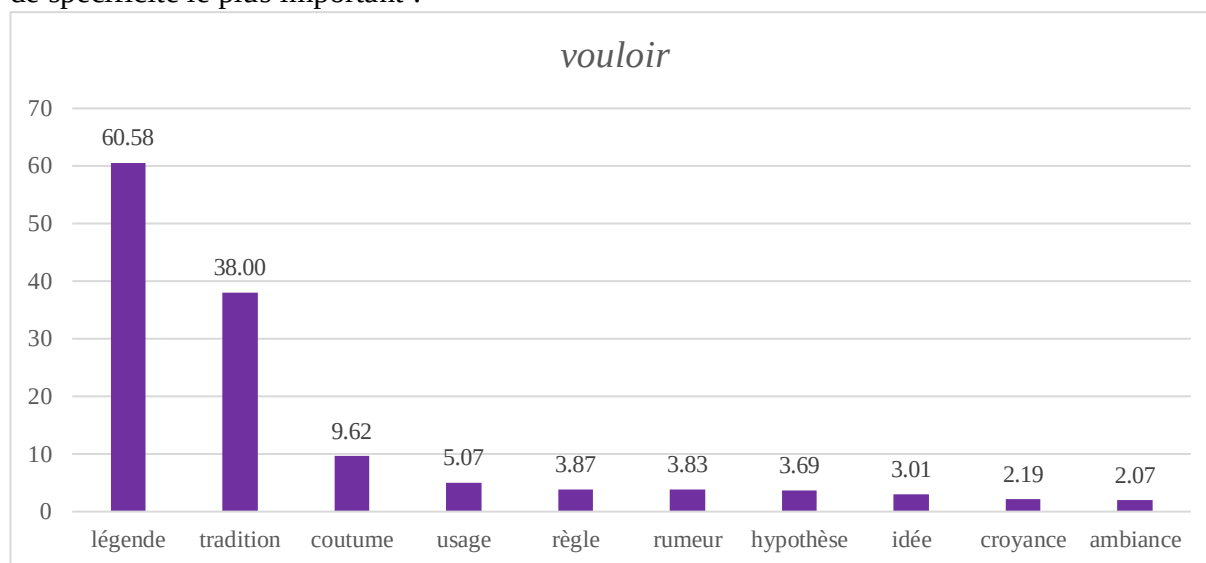


Figure 28 : Les 10 sujets nominaux de type objet cognitif pour *vouloir* sous sa forme lemmatisée dans le corpus WIKI selon l'indice de spécificité.

Avec ce type de nom, nous trouvons les mêmes variétés d'emploi que mentionnés avec *vouloir* au conditionnel dans LM :

132. Certains noms propres, par exemple ceux de pays, subissent cette assimilation, même si, surtout en volapük moderne, le critère de brièveté est moins respecté. Pour ceux qui n'ont pas de forme assimilée « canonique », il n'y a pas de règle absolue. **L'usage veut** que le mot soit gardé le plus possible proche de sa forme originelle, mais précédée par l'article *el* : c'est sur cet article que sera portée la déclinaison. (WIKI)
133. Selon **la tradition** qui **veut** que ce soit l'apprenti qui prenne la pose, Léonard aurait servi de modèle à la statue en bronze du « David » de Verrocchio. (WIKI)
134. La fontaine proche de la chapelle Notre-Dame-du-Vray-Secours [...] aurait eu des vertus curatives connues dès le Haut Moyen Âge. La **coutume voulait** qu'elle guérisse les enfants invalides ou en retard physiologiquement. (WIKI)
135. Revenu à Rome en octobre, César y célèbre son cinquième triomphe. César commet là une erreur politique que Plutarque soulignera : la **règle veut** qu'un triomphe honore une victoire sur un peuple ennemi de Rome, ce qui n'est pas le cas dans cette guerre civile. (WIKI)

Ces exemples font référence à des *normes* et à des *règles* comme susmentionné pour le corpus de presse (7.1.3.12). Cependant, ici, les tournures ne sont pas seulement importantes au conditionnel mais bien pour l'ensemble des formes de *vouloir*. On pourrait ainsi parler d'un *vouloir* qui aurait un usage « normatif », qui permet d'explicitier des *règles*, des *traditions* ou

des *logiques* et dont l'effet de sens est ainsi inféré à l'aide du co(n)texte. Une lecture à valeur modale déontique serait ainsi possible en prenant en compte l'inférence que les énoncés permettent ici. Nous pourrions ainsi admettre qu'il s'agit d'expression de la *volonté* d'une tierce personne ou d'une instance, mais inférée. Dans ce cas-là, l'*intention* ou la *volonté* serait encore perceptible : ce sont les *traditions*, *règles* ou encore *logiques voulues* par une personne ou un groupe qui sont exposées ici. Ces formulations semblent être une sorte de glissement métaphorique, qui peut être perçu dans les cas d'extension de sens – comme nous le mentionnions avec Recanati (2004) en 4.1. – mais qui peuvent aussi être propres au phénomène de grammaticalisation – comme l'explique Heine (2002) lorsqu'il énumère les facteurs propices à ce phénomène (4.1.1.). Dans ces cas, le NS de *volonté* ou d'*intention* devient moins facilement percevable.

Nous pourrions aussi accepter dans ces cas une lecture de valeur modale descriptive : en effet, dans ces exemples nous avons simplement une description d'un état de fait, d'une explication de ce qu'est la *tradition*, la *règle* ou encore la *logique* dans ces cas précis. Selon cette interprétation, le NS de l'*intention/volonté* est encore plus complexe à récupérer par inférence ce qui indiquerait un degré de grammaticalisation plus avancé.

Parmi ces sujets catégorisés comme des *objets cognitifs* nous pouvons aussi retrouver des noms dénotant davantage des concepts moins certains que des *normes*. C'est le cas des noms comme *idée*, *rumeur*, *hypothèse*, *légende*, ou encore *croyance* que nous avons classés comme des noms d'*a priori* précédemment (7.1.3.12) :

136. La Gaule, contrairement à l'**idée** préconçue qui **veut** qu'elle soit couverte de forêts dans lesquelles les Gaulois pratiquent essentiellement la chasse, était largement défrichée pour constituer des terres agricoles très riches avec de nombreuses fermes. (WIKI)
137. Une **rumeur** tardive **veut** qu'elle ait souffert de polydactylie (six doigts à sa main gauche) et d'une tache de naissance sur le cou, qu'elle cachait en tout temps. (WIKI)
138. Bien que les causes réelles restent à ce jour sujet à discussion, une **hypothèse** prédominante **veut** que cette forme d'asthme soit causée par une réaction auto-immune. (WIKI)
139. Une **légende** **veut** que les deux sœurs enterrées en ce lieu moururent de chagrin d'amour. (WIKI)
140. La **croyance** populaire **veut** que chaque personne mordue par un vampire finisse par devenir vampire à son tour. (WIKI)

Ces exemples illustrent un usage que l'on pourrait rapprocher des *suppositions*, donc des considérations d'ordre théorique qui ne sont pas vérifiées ou vérifiables. Ainsi, une valeur modale épistémique par inférence serait également possible avec ce type de sujet.

En somme, si nous résumons nos observations pour les associations de *vouloir* avec des *objets cognitifs*, nous avons des usages de *vouloir* normatif ou de *supposition* pas encore relevés dans la littérature mais très présents dans le corpus encyclopédique que nous employons dans cette recherche. Nous pouvons rappeler les réflexions de Dubois (1970) qui qualifiait les dictionnaires et les encyclopédies comme des discours pédagogiques. En ce sens, « le dictionnaire enseigne, ce qui signifie qu'il indique dans des instructions sans ambiguïté les questions qui peuvent être formulées et les réponses qu'en obtiendra le questionneur » (Dubois 1970 : 35). C'est ce genre d'emplois que *vouloir* semble véhiculer.

Nous identifions ici des effets de sens qui n'ont pas encore été discutés dans la littérature, avec *vouloir* précédé d'un sujet classé parmi les *objets cognitifs* selon la typologie de Salvadori et al. (2021). Cette construction semble instaurer une certaine distance par rapport aux *normes* ou aux *suppositions* mises en avant par la locutrice – puisque la ou les personnes désirant ou demandant ces normes ne sont pas mentionnées – ce qui rappelle également la réflexion que nous avons émise concernant le ménagement des faces selon la théorie de la politesse de Brown et Levinson (1987). L'utilisation de *vouloir* avec ces noms peut donc être considérée comme grammaticalisées à un stade plus ou moins avancé selon la valeur modale qu'on lui attribue.

7.1.4.5. Les sujets de type animé de devoir sous sa forme lemmatisée

Contrairement à son profil combinatoire dans LM (figure 14), *devoir* attire dans WIKI majoritairement des *animés* en position sujet. Non seulement il diffère ainsi de son comportement dans le corpus de presse, mais il se différencie également sur ce point de *pouvoir* dans WIKI. Nous allons donc observer de plus près quels sont les noms de cette catégorie qui ont l'indice de spécificité le plus important en position sujet pour *devoir* en tant que lemme :

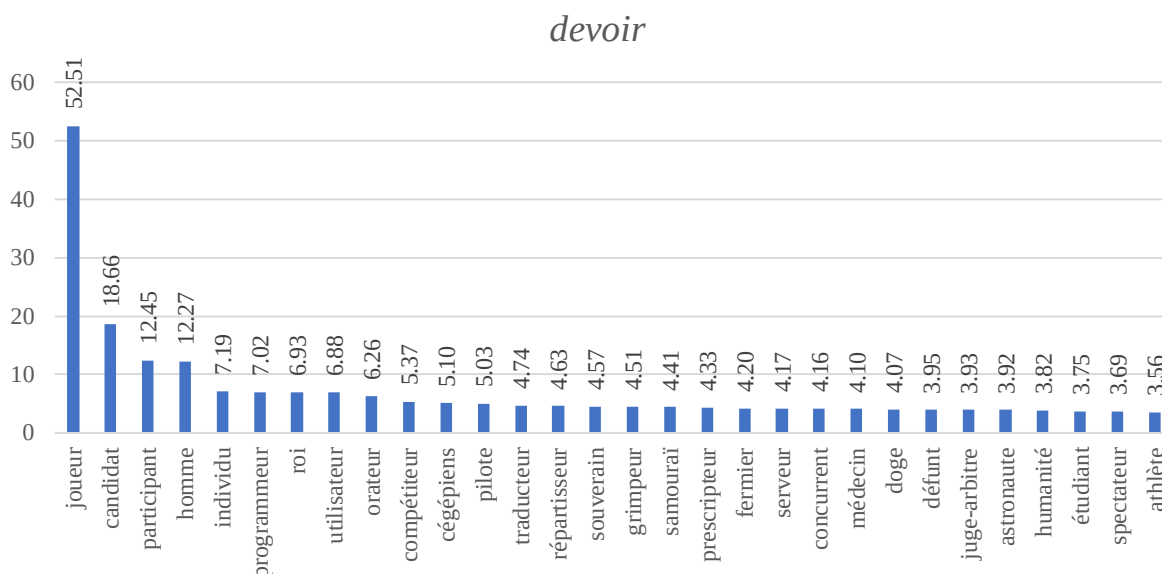


Figure 29 : Les 30 sujets nominaux de type animés avec le plus grand indice de spécificité pour *devoir* sous sa forme lemmatisée dans le corpus WIKI.

Avec cette liste, nous pouvons apercevoir que le texte exerce bien une influence sur le type de nom que les verbes attirent en position de sujet vu que nous n'avons pas ici des noms qui sont explicitement liés au domaine politico-économique comme nous l'avions dans LM. Cependant, concernant les effets de sens ou valeurs modales qu'on peut conférer à *devoir* dans ces configurations, il s'agit toujours de la modalité déontique, du moins, dans l'échantillon que nous avons analysé :

141. Léonard de Vinci pense que l'**homme doit** s'engager activement à combattre le mal et faire le bien, car « Celui qui néglige de punir le mal aide à sa réalisation ». (WIKI)
142. Pour se présenter, un **candidat doit** être électeur. (WIKI)
143. L'escalade est un jeu de (dé) placements et d'équilibre. Le **grimpeur doit** apprendre à progresser et à gérer son centre de gravité dans un univers vertical, et acquérir ainsi un vocabulaire gestuel. (WIKI)
144. [...] l'Apple I était constitué uniquement d'une carte assemblée comprenant des composants électroniques dont environ 21 circuits intégrés. Cependant, pour en faire un ordinateur fonctionnel, les **utilisateurs devaient** encore l'intégrer dans un boîtier avec une alimentation, un clavier, et un écran de télévision. (WIKI)
145. Pour bénéficier de l'aide d'Anubis — fils et protecteur d'Osiris — le **défunt doit** démontrer qu'il est digne d'être perçu comme un autre Osiris. (WIKI)
146. L'objectif du joueur est de traverser trois séries de niveaux, chaque niveau du jeu étant peuplé de créatures hostiles que le **joueur doit** combattre, à l'aide d'armes qu'il récupère au fur et à mesure de sa progression, pour pouvoir avancer et en atteindre la sortie. (WIKI)

147. En 2025, dans des États-Unis devenus une dictature, un jeu fait fureur : "La Grande Traque". Un **participant doit**, durant trente jours, échapper aux tueurs qui le poursuivent, tueurs aidés par la population qui est encouragée à donner sa localisation. (WIKI)
148. Le véritable **traducteur doit** être capable de se mettre dans la peau d'un écrivain afin de capter et de saisir l'intention (le « vouloir dire ») de l'auteur du texte de départ. (WIKI)

Nous pouvons constater que, au sein de l'encyclopédie WIKI, l'effet de sens d'*obligation* attribué au verbe *devoir* fonctionne comme une *prescription*. Cet effet de sens est inféré notamment grâce aux prémices que nous pouvons avoir en tant que lectrice d'un texte encyclopédique. En effet, l'encyclopédie semble détailler les règles et les lois régissant certains aspects de la vie et du monde en utilisant le verbe *devoir*. Ce phénomène pourrait également résulter du lexique statistiquement associé à *devoir*, pouvant faire référence à des *modes d'emploi* ou des *règles de jeu*, tels que *joueur*, *candidat*, *participant*, *utilisateur*, *compétiteur*, *concurrent*, voire *juge-arbitre*. Ce type d'utilisation peut rapprocher *devoir* de *vouloir* avec des sujets grammaticaux de type *objet cognitif*, qui dans ces constructions permet d'explicitier des normes tout en maintenant une certaine distance grâce à l'usage du modal. En outre, ce type d'utilisation peut également être assimilé à *pouvoir*, lorsqu'il exprime des généralités avec la valeur modale descriptive.

7.1.4.6. La granularité épaisse pour les formes au présent de l'indicatif

La granularité épaisse avec les formes au présent de l'indicatif nous permet de voir si les observations que nous avons soulevées pour les formes lemmatisées s'appliquent pour ce temps verbal ou si celui-ci permet de faire ressortir d'autres propriétés sémantiques pour ces verbes modaux par rapport à leurs profils combinatoires :

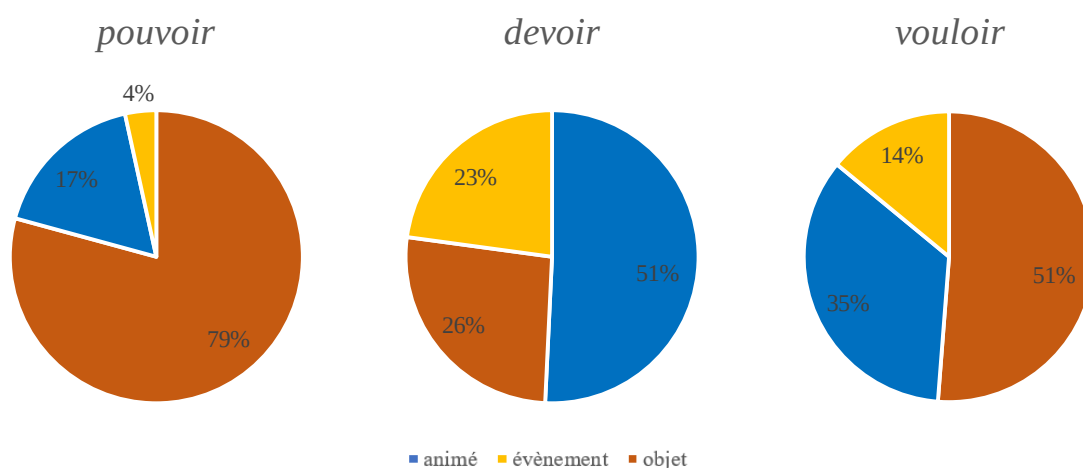


Figure 30 : Répartition des noms en position sujet pour les formes à l'indicatif présent des modaux selon la granularité de Varvara et al. (sous presse) dans le corpus WIKI.

Les trois verbes présentent des divergences significatives entre eux, ainsi qu'avec les formes lemmatisées précédemment analysées dans ce même corpus.

Dans la figure 32, le verbe *pouvoir* se distingue par une préférence nette pour les *objets*. *Devoir* manifeste une préférence pour les sujets animés, contrairement aux résultats disparates obtenus pour les formes lemmatisées. En ce qui concerne *vouloir*, les *objets*, à l'indicatif présent, sont les noms privilégiés en position de sujet, une tendance qui n'était pas observée dans les formes lemmatisées de ce corpus ni dans l'ensemble des profils combinatoires étudiés dans le corpus LM. Il s'agit donc d'une préférence particulière de *vouloir* au présent de l'indicatif dans WIKI. Nous allons désormais voir avec la granularité moyenne si ce sont les mêmes types d'*objets* qui apparaissent de manière privilégiée avec *pouvoir* et *vouloir* mais aussi si d'autres sous-catégories apparaissent comme importante.

7.1.4.7. La granularité moyenne pour les formes à l'indicatif présent

La granularité moyenne de la typologie de Salvadori et al. (2021) confirme ce que nous avons observé pour les formes lemmatisées dans ce même corpus :

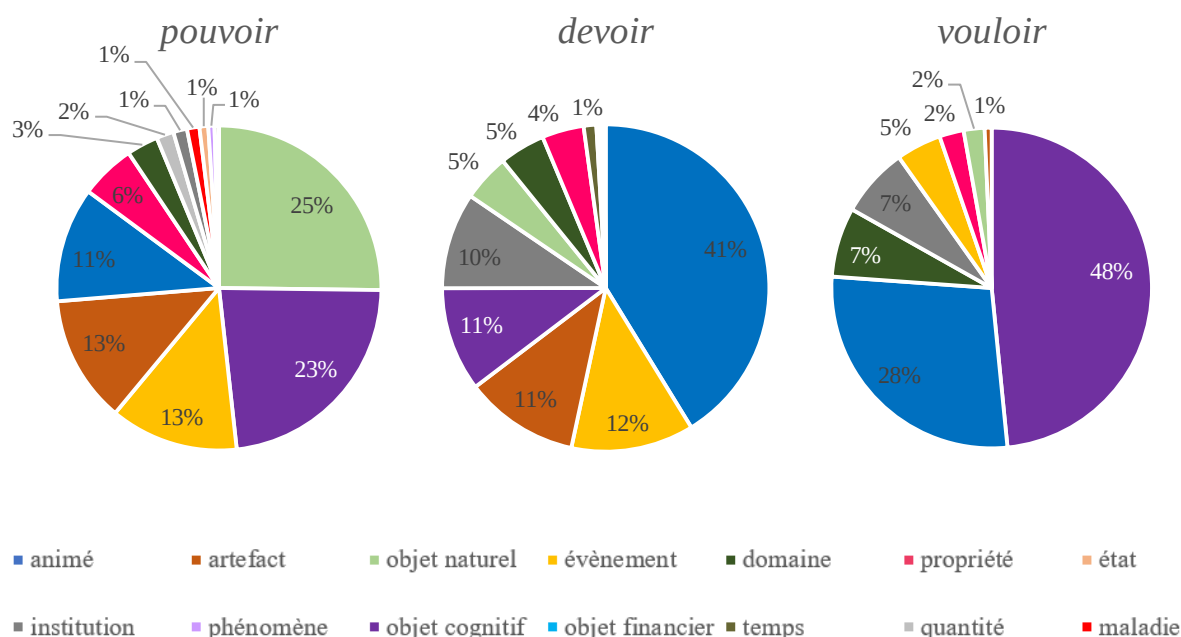


Figure 31 : Répartition des noms en position sujet pour les formes au conditionnel présent des modaux selon les catégories de Salvadori et al. (2021) dans le corpus WIKI.

Ainsi, pour *pouvoir*, les *objets naturels* et *cognitifs* sont représentés de manière presque égale lorsque le verbe est à l'indicatif présent, ce qui ne diverge pas des observations pour les formes lemmatisées. Cependant, les *animés* sont ici moins importants proportionnellement parlant. Pour *devoir*, les *animés* sont ici représentés presque une fois et demie de plus que pour la forme lemmatisée. Cependant, l'ordre d'importance des catégories sémantiques reste le même que ce soit sous sa forme lemmatisée ou au présent de l'indicatif. Enfin, c'est pour *vouloir* que la démarcation semble la plus importante par rapport aux deux autres verbes mais aussi à l'égard de ce que nous avons déjà pu admettre à son sujet. Il semble ainsi se grammaticaliser dans ce corpus au présent de l'indicatif.

Au vu de la similarité des propriétés sémantiques les plus importantes des formes du présent de l'indicatif avec les formes lemmatisées, nous n'allons pas entrer ici en détail dans les indices de spécificité.

7.1.4.8. La granularité épaisse pour les formes au conditionnel présent

Les résultats quant à la répartition des sujets de type nominal pour les formes au conditionnel présent montrent des propriétés à contraster avec les résultats pour le présent de l'indicatif ou avec les formes lemmatisées :

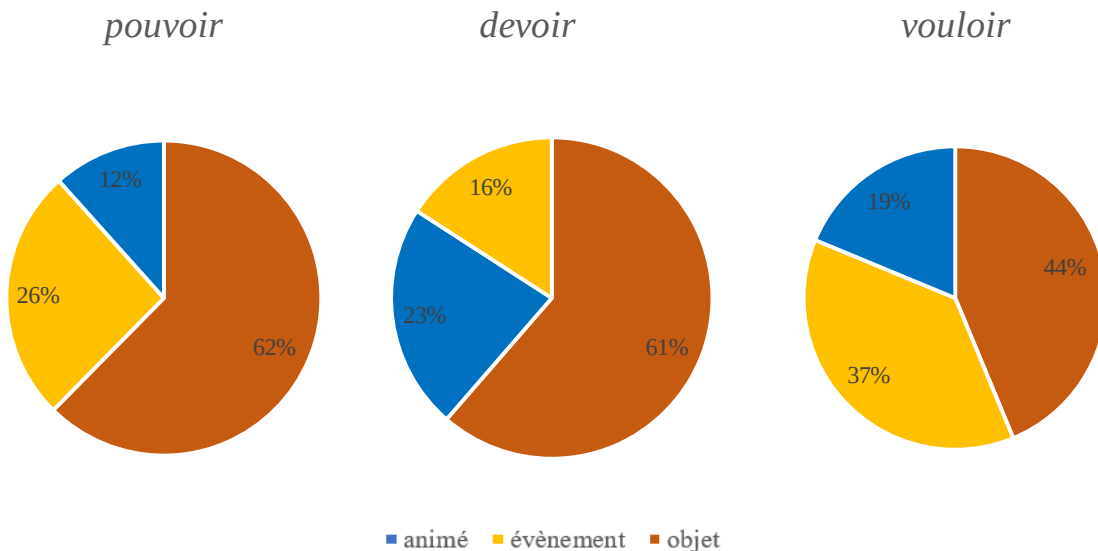


Figure 32 : Répartition des noms en position sujet pour les formes au présent de l'indicatif des modaux selon la granularité de Varvara et al. (sous presse) dans le corpus WIKI.

Le profil combinatoire de *pouvoir* au conditionnel présent ressemble ici beaucoup à celui sous sa forme lemmatisée dans ce même corpus. Par contre, en ce qui concerne *devoir* et *vouloir* ceux-ci diffèrent par rapport aux observations soulevées pour les autres paramètres étudiés.

Ainsi, *devoir* attire ici majoritairement des *objets* en position sujet dans la même proportion que *pouvoir*. Quant à *vouloir*, c'est la première fois depuis le début de nos observations que nous retrouvons les sujets *animés* en minorité dans son profil combinatoire.

7.1.4.9. La granularité moyenne pour les formes au conditionnel présent

La granularité moyenne nous permet ici de confirmer la ressemblance entre *pouvoir* et *devoir* mais aussi d'apercevoir que ce sont à nouveau majoritairement les *objets cognitifs* qui sont les sujets les plus importants pour *vouloir* dans ce corpus :

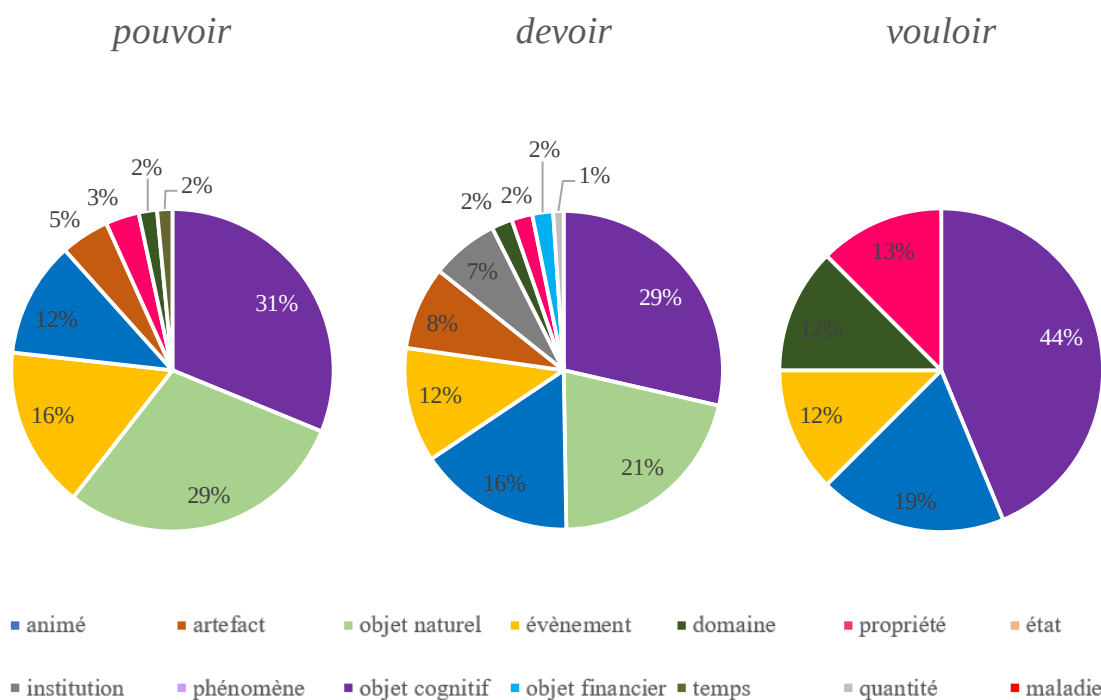


Figure 33 : Répartition des noms en position sujet pour les formes au conditionnel présent des modaux selon les catégories de Salvadori et al. (2021) dans le corpus WIKI.

Pouvoir et *devoir* ont tous deux les *objets cognitifs* puis les *objets naturels* en première et deuxième position des catégories les plus représentées parmi les sujets nominaux avec un indice de spécificité significatif dans WIKI lorsqu'ils sont au conditionnel présent. Ainsi, les trois verbes ont, dans cette configuration, une majorité d'*objets cognitifs* en sujet.

Nous allons donc observer de plus près les occurrences de *pouvoir* et *devoir* au conditionnel avec ces types d'objet tout comme *vouloir* avec les *objets cognitifs*.

7.1.4.10. Les sujets de type objet naturel et cognitif de pouvoir et devoir au conditionnel présent

Pouvoir et *devoir* attirent tous deux en majorité des *objets naturels* et *cognitifs* lorsqu'ils sont conjugués au conditionnel présent. Cependant, en étudiant les noms qui possèdent les indices de spécificité les plus importants dans ces catégories pour les deux verbes, nous pouvons noter qu'aucun nom n'est ici en commun :

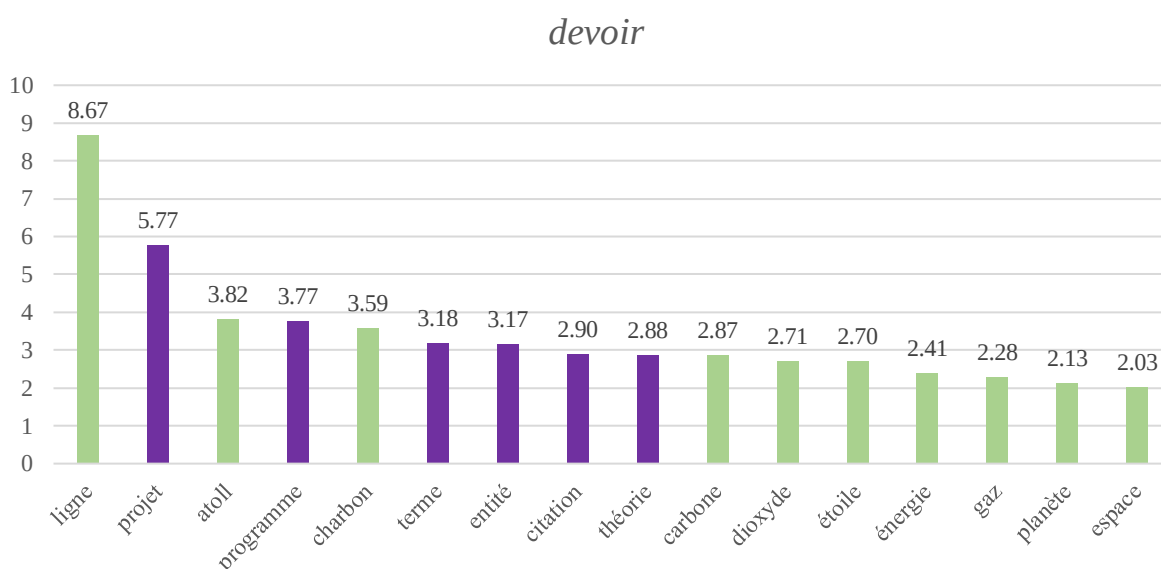
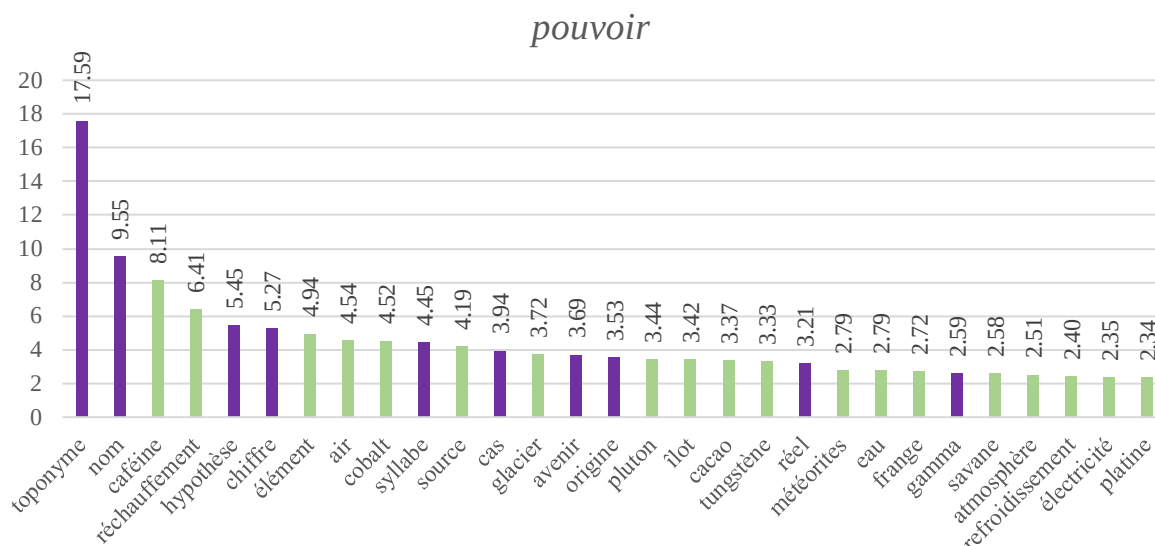


Figure 34 : Les 30 sujets nominaux de type objet naturel et cognitif avec le plus grand indice de spécificité pour pouvoir et les 17 sujets nominaux de type objets naturels et cognitifs avec un indice de spécificité significatif pour devoir au conditionnel présent dans le corpus WIKI.

Même si aucun nom n'est partagé, il y a tout de même des regroupements possibles : ainsi, *pouvoir* attire le nom *hypothèse* en position sujet et *devoir* le terme *théorie*, ce qui rappelle les sujets significatifs de *vouloir* sous leur forme lemmatisée que nous avons définis comme étant des noms dénotant des *a priori*. Nous retrouvons ici un effet de sens de *supposition* comme nous le mentionnions pour *vouloir* avec ces associations :

149. Toutefois, le choix du premier homme à marcher sur la Lune fut remis en question du fait de la prééminence naturelle du commandant et de l'agencement du module (l'écouille s'ouvrait de l'intérieur vers la droite, ce qui facilitait la sortie pour Armstrong et la compliquait pour Aldrin). [...] Une autre **hypothèse pourrait** expliquer le choix

d'Armstrong : celui-ci était civil (NASA) et non militaire (USAF) comme ses deux compagnons. Armstrong semblait effectivement un bon choix : *Mister Cool* comme le surnommaient ses collègues, était « réputé pour son humour décalé mais surtout son sang-froid, son calme [et] sa capacité à prendre la bonne décision ». (WIKI)

150. Il n'existe pas actuellement une théorie de la gravité quantique satisfaisante du point de vue de la phénoménologie, bien que la théorie des supercordes soit un bon candidat (la gravitation quantique à boucles cependant ne propose pas d'unifier la gravitation avec les autres interactions du modèle standard). En revanche, une **théorie** quantique gravitationnelle **devrait** posséder les caractéristiques suivantes [...] (WIKI)

Toutefois, cette interprétation n'est pas systématique, contrairement aux occurrences que nous avons relevées de ce type de nom avec *vouloir*. *Pouvoir* peut très bien dénoter la *possibilité* (151) et *devoir* un fait futur (152) :

151. Le retour moderne de la Terre-mère pourrait signer l'apparition de nouveaux mythes selon certaines personnes. Par exemple, le mythologue Joseph Campbell pensait que l'**hypothèse** Gaïa **pourrait** être un futur mythe, qui parlerait non pas d'une localité ou d'un peuple, mais d'une planète entière, avec tous les êtres vivants qui s'y trouvent. Le nouveau mythe indiquerait comment entrer en rapport avec la nature et le cosmos, et la société concernée par le mythe serait une société planétaire. (WIKI)
152. Einstein s'aperçut alors que cette propriété du rayonnement était en opposition de manière irréductible avec la théorie électromagnétique classique (élaborée par Maxwell). Dès 1906, il annonça que cette **théorie devrait** être modifiée dans le domaine atomique. La manière dont cette modification devrait être obtenue n'était pas évidente puisque la physique théorique reposait sur l'utilisation d'équations différentielles, dites équations de Maxwell, correspondant à des grandeurs à variation continue. (WIKI)

152 peut également, par inférence, être interprété comme un ordre : en effet, Einstein – ayant une certaine position d'autorité dans ce contexte – *demande* à ce que la théorie soit modifiée. Concernant les noms qui sont classés parmi les *objets naturels*, nous trouvons aussi des effets de sens très variés pour ces deux verbes. Nous pouvons par exemple repérer des occurrences où *pouvoir* et *devoir* permettent d'exprimer des *prédictions* :

153. En 2019, l'EIA estime que l'**électricité** nucléaire **pourrait** aux Etats-Unis chuter de 17 % en 2025 par rapport au niveau de 2018 [...] (WIKI)
154. Au rythme actuel, les **glaciers pourraient** disparaître d'ici 200 à 300 ans. (WIKI)

Lorsque les *prédictions* sont moins certaines – car non-datées – nous pouvons alors plutôt parler d'*hypothèses* comme dans les exemples suivants :

155. Ces **éléments pourraient** avoir des isotopes présentant des périodes radioactives atteignant quelques minutes. (WIKI)
156. De l'évolution de cette courbe de luminosité, une mince atmosphère de 0,15 Pa a été déterminée, soit environ un 700 000^e de celle de la Terre. Cette **atmosphère pourrait** n'exister que lorsque la planète est proche de son périhélie, et geler lorsqu'elle s'éloigne du Soleil. (WIKI)

Cet effet de sens d'*hypothèse* peut également concerner des faits que l'on déduit à partir d'éléments du passé :

157. À l'exception de celui de Christmas, qui est le plus ancien et le plus grand atoll au monde, ces **atolls ne devraient** avoir complètement émergé, à partir de « makatea », qu'au tout début de l'ère chrétienne. (WIKI)
158. Le premier **élément** de Garonne **pourrait** provenir soit d'un radical pré-latin " * gar- " « pierre, rocher, montagne », variante de " * kar- " et de " * gal- ", soit, plus vraisemblablement, d'un radical bilabial *gw-ar- qui expliquerait mieux les différentes variantes observées. (WIKI)
159. D'après Godefroid Kurth, « il n'y manque que l'élément merveilleux pour le classer dans la catégorie des histoires en l'honneur des saints ». Mais l'historien belge, ne reconnaissant cependant dans ce récit aucun des critères qui lui semblaient signer une origine populaire, accepte son historicité : il suggérerait même que la **source** originale **pourrait** être la *Vita* perdue de saint Remi que Grégoire affirme par ailleurs avoir eue sous les yeux et que le récit pourrait ainsi remonter à un témoin proche et peut-être oculaire de l'évènement (WIKI)
160. La ferme des Boves est une vaste demeure dont certains **éléments pourraient** dater de la fin du XVI^e siècle. (WIKI)
161. Le bâtiment est d'inspiration classique dont les **sources pourraient** se chercher de façon directe ou indirecte chez l'architecte romain Vitruve ou encore chez Andrea Palladio, architecte italien de la Renaissance qui eut une influence considérable sur toute l'architecture occidentale. (WIKI)

La certitude semble plus importante lorsque *devoir* est employé dans des co(n)textes similaires. Dans les exemples suivants, compte tenu des connaissances, nous pouvons alors parler de *quasi-certitude* et donc de *prédiction* sur la façon dont les choses *devraient* se dérouler :

162. Bode a fait valoir que, tout comme Saturne était le père de Jupiter, la nouvelle **planète devrait** être nommée d'après le père de Saturne. (WIKI)
163. Une définition alternative considère que les **planètes devraient** être distinguées des naines brunes sur la base de leur formation. (WIKI)
164. Encore sur la question de l'eau de pluie, plus précisément pour les systèmes de récupération des eaux pluviales pour la consommation, l'arrosage des légumes ou la consommation des animaux, il a été suggéré de tamponner l'acidité de la pluie et de la filtrer sur charbon actif. Ce **charbon devrait** ensuite être brûlé dans des incinérateurs équipés de filtres appropriés. (WIKI)

A la lumière des exemples que nous avons relevés dans cette partie, nous pouvons admettre qu'il existe des nuances fines d'effets de sens acceptables pour ces deux verbes lorsqu'ils sont précédés par des *objets naturels* ou *cognitifs* mais que ces effets de sens peuvent se rapprocher, pour la plupart de ces exemples, d'une valeur modale épistémique vu qu'une certaine incertitude règne dans chacun de ces énoncés. A nouveau, nous ne pensons pas que cela soit dû aux sujets grammaticaux en tant que tels, mais plutôt au mode du conditionnel.

7.1.4.11. Les sujets de type objet cognitif de vouloir au conditionnel présent

Comme mentionné, les *objets cognitifs* représentent la plus grande part des sujets grammaticaux avec un indice de spécificité significatif pour *vouloir* au conditionnel dans WIKI. Cependant, il est important de nuancer que cela représente seulement trois noms en tant que tels :

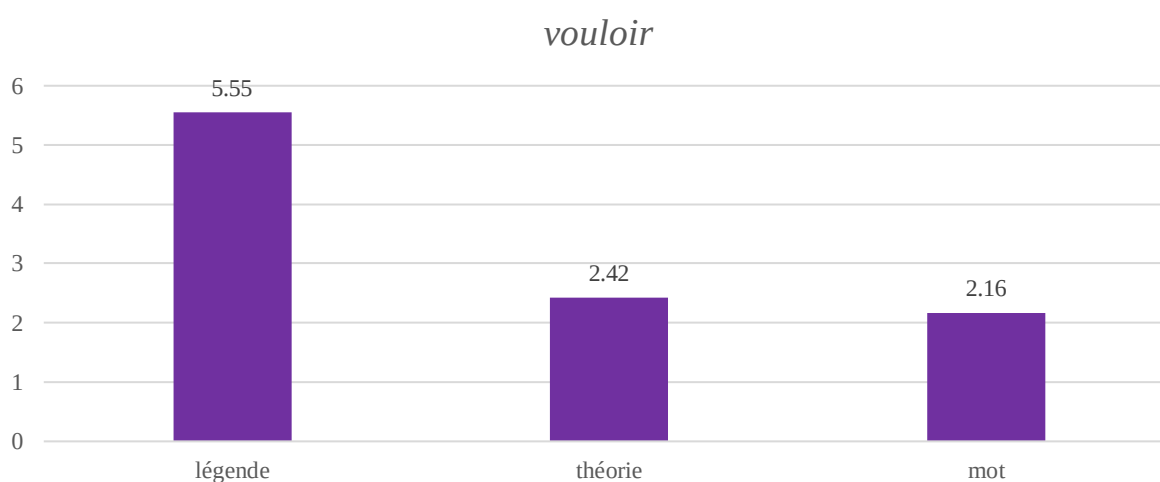


Figure 35 : Les 3 sujets nominaux de type objets naturels et cognitifs avec un indice de spécificité significatif pour *vouloir* au conditionnel présent dans le corpus WIKI.

Comme pour sa forme lemmatisée, nous avons ici des noms qui dénotent des *suppositions* comme *légende* et *théorie* :

165. Une autre **théorie voudrait** que le nom de Saint-Thonan provienne de saint Donan, moine irlandais, mais cette hypothèse semble à rejeter, ce saint étant un disciple de Briec avec Fragan [...] (WIKI)
166. La **légende voudrait** que l'analyse dimensionnelle ait permis à Geoffrey Ingram Taylor d'estimer en 1950 l'énergie qu'avait dégagée l'explosion d'une bombe atomique, alors que cette information était classée top secret. (WIKI)
167. Une autre **théorie voudrait** que « protéine » fasse référence, comme l'adjectif protéiforme, au dieu grec Protée qui pouvait changer de forme à volonté. Les protéines adoptent en effet de multiples formes et assurent de multiples fonctions. Toutefois, cette caractéristique ne fut mise en évidence bien plus tard, au cours du XX^e siècle. (WIKI)
168. Une **légende voudrait** que ce soit Marco Polo qui ait rapporté [...] les pâtes de Chine. Celle-ci est en réalité infondée. En réalité, les Grecs et les Romains connaissaient déjà des pâtes fraîches du type lasagnes (« laganon » en grec, « laganæ » en latin). (WIKI)

Contrairement à ce que nous observions dans la partie sur les formes lemmatisées – dont les exemples étaient pour la plupart au présent – il se trouve que ces *suppositions* sont infondées avec *vouloir* au conditionnel.

Concernant les occurrences avec le sujet *mot*, nous pouvons apercevoir qu'il s'agit de *supposition* ou d'*hypothèse* sur l'origine de ces mots. Le verbe *vouloir* est ici suivi du verbe *dire*, ce qui montre ici que nous sommes bien face à une explication d'une *signification* :

169. La forme première du nom de Bréhal est « Brohal » (« Bro » en langue bretonne veut dire pays et « hal », le sel). En langue celte, « Bré » signifie colline. Le **mot** Bréhal **voudrait** donc dire « pays de collines » ou « pays de sel ou de mer ». (WIKI)
170. Une histoire raconte que le nom « Afghan » vient du **mot** qui **voudrait** dire « cavalier ». Les gens du peuple, pour faciliter la prononciation, disaient « Apagan ». La phonétique changea lors de la venue des Arabes. Dans l'alphabet arabe la lettre « p » n'existait pas alors ce qui donna « Afagan ». Ce mot évolua pour enfin donner le mot Afghan. (WIKI)

Franckel (2018) qualifie *vouloir dire* de marqueur de *reformulation*. Or, il s'avère qu'au conditionnel, il y aurait plutôt *supposition* d'une *signification*.

En somme, les occurrences de *vouloir* au conditionnel présent avec les sujets que nous avons classés comme étant des *objets cognitifs* confirment sa grammaticalisation dans ce type de construction.

7.1.5. Le cas de *falloir*

Il peut être étonnant de trouver *falloir* dans l'étude des sujets grammaticaux des verbes modaux étant donné qu'il est, par essence, un verbe impersonnel. Nous avons pris le parti de le considérer dans notre étude car il arrive qu'un sujet nominal soit exprimé pour *falloir* dans la complétive. C'est le cas notamment dans l'exemple suivant :

171. Il cherche à créer un groupe uni et solidaire, qui ait une identité forte. Pour cela, **il faut** que **cette collectivité** partage une histoire et un destin communs et qu'elle se construise sur la volonté de perpétuer son climat actuel. (WIKI)¹⁶³

Dans ces cas, nous parlons ainsi de *sujet sémantique* et non pas *syntactique* contrairement aux autres verbes que nous avons analysés ici. Cette différenciation a été documentée notamment par Girard (2004) pour les phrases passives mais elle peut s'appliquer à l'analyse de tournures impersonnelles comme c'est le cas avec *falloir*. Ainsi, dans l'exemple 171, le sujet syntactique est bien celui avec lequel le verbe s'accorde (pour *falloir*, le pronom impersonnel *il*) mais le sujet sémantique est *cette collectivité* qui est dans la *contrainte* de partager une histoire et un destin, et qui est, en somme, le sujet de la complétive. Dans ces cas, comme pour *pouvoir* et *devoir*, le sujet exprimé est ainsi en lien avec le verbe modalisé.

Lorsque le sujet est exprimé, il peut donner des indications supplémentaires concernant les caractéristiques sémantiques de *falloir*. Cependant, contrairement à *pouvoir*, *devoir* et *vouloir*, il n'y a pas de recherche existante qui se concentre sur cet élément pour interpréter sémantiquement *falloir*. Nous ne pouvons ainsi pas comparer nos résultats et nos interprétations à celles d'autres autrices, mais seulement proposer de nouvelles pistes pour appréhender les caractéristiques sémantiques de ce verbe à l'aide de cet élément du cotexte privilégié.

De par sa nature de verbe impersonnel, *falloir* comporte bien moins de sujets nominaux que les autres verbes que nous étudions ici. Ainsi, ces sujets exprimés représentent 5% des occurrences de *falloir* dans sa forme lemmatisée dans LM et 3% seulement dans WIKI :

¹⁶³ Afin de chercher les sujets nominaux qui apparaissent dans la complétive de *falloir* en tant que lemme, nous avons saisi la requête suivante dans TXM : [frlemma="falloir"][frpos!="V.*"]?[frpos!="V.*"]?[frpos!="V.*"]?[frpos!="V.*"]?[frpos!="V.*"]?[frpos!="V.*"] dans un empan de 2 après le pivot. Cette requête permet d'éviter qu'un verbe à l'infinitif apparaisse entre *falloir* et la complétive. En effet, si c'était le cas, le sujet exprimé dans la complétive ne dépendrait ainsi plus de *falloir* mais du verbe de la complétive comme dans « Il ne **faut** jamais **oublier** que **le théâtre est un rituel** » (LM, 15.04.2010). Ici, c'est le verbe *oublier* qui est modalisé par *falloir* et non pas la complétive qui dépend d'*oublier*.

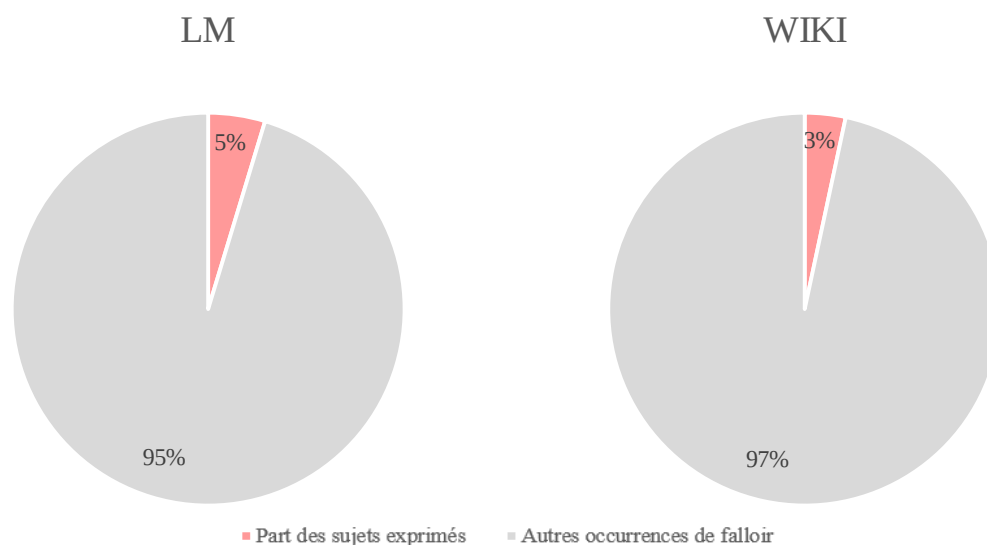


Figure 36 : Part des sujets nominaux exprimés dans la complétive pour falloir sous sa forme lemmatisée.

En nous intéressant aux cooccurrences spécifiques où le sujet de *falloir* est exprimé, nous pouvons observer les cotextes dans lesquels la locutrice estime qu'il est indispensable de signaler sur qui ou quoi porte la *contrainte* ou la *nécessité*, habituellement générique avec la nature impersonnelle de ce verbe¹⁶⁴.

Pour le corpus de presse LM, les sujets exprimés dans la complétive sont relativement diversifiés mais les *animés* représentent tout de même la majorité des noms dans cette fonction :

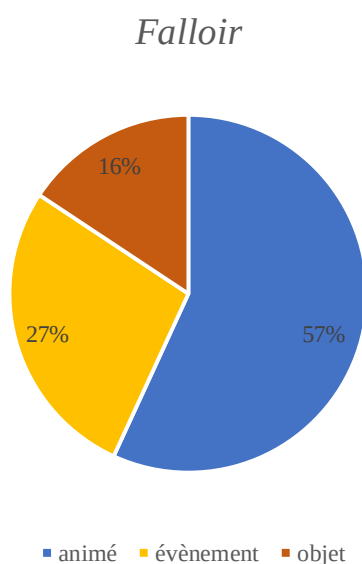


Figure 37 : Répartition des noms en position sujet selon la granularité de Varvara et al. (sous presse) pour falloir sous sa forme lemmatisée dans le corpus LM.

¹⁶⁴ En cela, nous pouvons également rappeler les remarques de Brown et Levinson (1987) dans leur théorie sur la politesse. Il est signalé dans leur ouvrage que les formes impersonnelles permettent de ménager la face (Brown & Levinson 1987 : 70). Il peut ainsi être intrigant de voir dans quels cas cette forme impersonnelle de *falloir* signale sur qui ou quoi porte la *nécessité*.

De manière plus détaillée, nous remarquons que ce sont des *animés* purs qui occupent la grande majorité de ces noms avec un indice de spécificité positif en position de sujet sémantique :

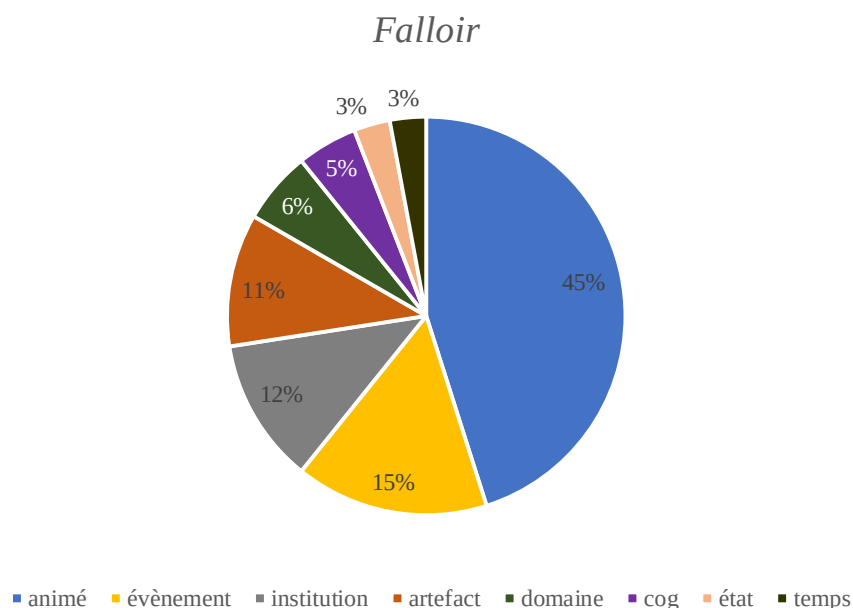


Figure 38 : Répartition des noms en position sujet pour les formes lemmatisées de falloir selon les catégories de Salvadori et al. (2021) dans le corpus WIKI.

Les noms en position sujet avec un indice de spécificité significatif pour *falloir* sous sa forme lemmatisée sont au nombre de 24. Même si les *animés* sont les plus représentés en termes de cooccurrences brutes, c'est avec un nom de type *artefact* que l'indice est le plus important mais qui représente une chose très générale ce qui permet finalement à *falloir* de garder ses propriétés de verbe impersonnel :

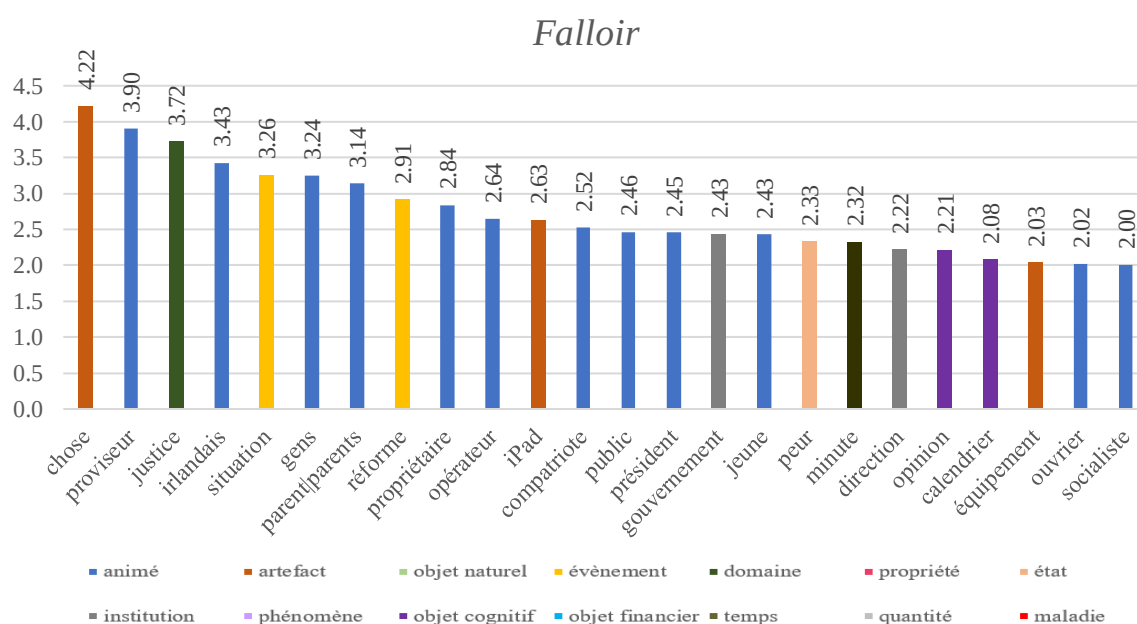


Figure 39 : Les 24 sujets nominaux avec un indice de spécificité significatif pour falloir sous sa forme lemmatisée dans le corpus LM.

Les valeurs modales déontiques restent bien présentes avec les sujets de type *animé* :

172. Pour être élu au Soudan sous le regard international, après vingt et un ans d'un pouvoir arraché à l'origine lors d'un coup d'Etat, il **a fallu** que le **président** Omar Al-Bachir renonce à la couleur verte. Celle de l'armée soudanaise et de son uniforme, une teinte qu'il décrivait comme sa « couleur politique » lorsqu'il avait été propulsé à la tête de la junte qui venait de prendre le contrôle à Khartoum, le 30 juin 1989. (LM, 30.04.2010)
173. Ainsi, le Nu au plateau de sculpteur de Picasso, adjugé 106, 5 millions de dollars frais inclus, le 4 mai, chez Christie's. Un record du monde qui a failli échoir à la concurrence : selon de bonnes sources, Christie's n'était pas la mieux placée pour l'emporter, les administrateurs de la succession préférant la maison rivale Sotheby's. Il **a** donc **fallu** que le **propriétaire** de la première, François Pinault, mette la main au portefeuille, et propose une garantie, c'est-à-dire une somme forfaitaire allouée au vendeur quel que soit le résultat des enchères. (LM, 04.06.2010)
174. Pour Marie-Pierre Vieu, membre de la direction, « ce départ n'était pas inéluctable. Il **aurait fallu** que la **direction** donne des signes pour les retenir ». (LM, 09.06.2010)
175. Si l'on veut que La Marseillaise soit chantée, il **faut** que les **jeunes** la connaissent et en aient l'habitude. (LM, 05.07.2010)
176. « L'important, c'est que cette aide soit mise en œuvre. Naturellement, il **faut** que les **réformes** (...) se réalisent, mais il y a des discussions tout à fait possibles sur le rythme, le contenu exact de ces réformes » (LM, 25.01.2010)

Les valeurs modales déontiques ressortent ici bien puisque nous pouvons remplacer *falloir* dans toutes ces phrases par *il est/a été nécessaire que*. La plupart de ces exemples possèdent un sujet *animé* ou de l'ordre d'une *institution* hormis l'exemple 176 avec le nom *réformes* que nous avons catégorisé dans le groupe des noms d'évènement mais dont la valeur modale déontique peut à nouveau être inférée car nous comprenons qu'une personne ou un groupe de personnes doit être à l'origine de ces réformes.

Très proches de ces exemples, mais avec une nuance fine, figurent les exemples suivants dans lesquels *falloir* permet d'exprimer une *condition nécessaire* au déroulement d'une autre action :

177. La rigueur est nécessaire et même incontournable pour sauver la Grèce, qui est en grand péril, mais il **faut** que la **justice** sociale soit au rendez-vous. Si elle ne l'est pas, ça devient insupportable pour les peuples. (LM, 07.05.2010)
178. Les Sénégalais évitent de parler d'« accord » de défense. Ils souhaitent que Paris sollicite l'utilisation des « facilités » de Dakar. Pareille demande de reddition symbolique est mal

- perçue à Paris, qui exige des assurances sur le statut juridique des militaires entre la cérémonie et leur départ effectif, dans au moins un an. Il **faudra** que le **président** sénégalais en personne, accompagné de son fils et de son ministre de la défense, soient reçus à Paris, fin mai, par le général Benoît Puga, chef d'état-major particulier du président Sarkozy, pour apaiser la méfiance réciproque et trouver un accord sur le principe d'une cérémonie de « restitution » de la base française, le 9 juin. (LM, 01.06.2010)
179. Le 31 mai, le tribunal administratif de Nice avait décidé que cette fermeture n'était pas justifiée. Il **a fallu** que la **direction** des musées des Alpes-Maritimes assure à ses frais le gardiennage de l'exposition pour que la décision du tribunal puisse être appliquée. (LM, 12.07.2010)
180. Il **faudrait** donc que la **situation** sociale s'aggrave nettement pour que l'ouverture cédétiste soit prise en compte. (LM, 07.09.2010)
181. M. Ayrault comme Mme Aubry ont rappelé un certain nombre de « préalables ». « Il **faut** que le **gouvernement** renonce à la procédure d'urgence, qu'il tienne compte des avis du Conseil d'Etat et que le texte soit applicable concrètement. Si ces préalables sont remplis, notre attitude sera ouverte », a indiqué le député de Loire-Atlantique. (LM, 29.04.2010)
182. Aujourd'hui le marché n'est pas ouvert. Il le sera au moment où l'Arjel aura délivré les premiers agréments. Cela pourrait être pendant la semaine du 7 juin ou juste avant. Tout dépendra de la qualité des dossiers. Il **faudra** aussi que les **opérateurs** aient déclaré opérationnel leur dispositif de captage et d'archivage des données et que leurs logiciels de jeux soient agréés. Mais ils ont six mois pour que leur système informatique soit conforme à leur déclaration. (LM, 20.05.2010)

Ces usages rappellent ce que Caron et Caron-Pargue (2003) avait nommé une *nécessité tournée vers un but* (4.4.3.). Cette lecture est récupérable grâce à la préposition *pour* en 177 et 178 et la conjonction *pour que* dans 179 et 180. Pour 181, c'est le nom *préalables* de l'énoncé qui précède qui nous permet d'inférer la conditionnalité de *falloir* qui apparaît dans la phrase qui suit. Pour 182, c'est l'inférence du cotexte en général qui nous aide à comprendre qu'il y a ici une condition supplémentaire à l'*ouverture des marchés* qui est exprimée avec *falloir* au futur et le sujet sémantique *opérateurs*.

En somme, il semblerait que ce ne soit pas tant le sujet exprimé dans la complétive qui permet une interprétation des caractéristiques sémantiques de *falloir* mais le cotexte plus global qui justifie une inférence plus ou moins aisée de ce que le verbe exprime.

Pour WIKI, seulement deux noms figurent en position de *sujet sémantique* de manière spécifique, il s'agit d'*enfant* et *président*, des noms *animés* :

183. Mais le FBI s'intéressa aussi, dès cette période, aux activistes politiques, mettant en œuvre une surveillance des mouvements politiques les plus divers. La « *red scare* » (« peur rouge ») affectait en effet les États-Unis à ce moment. Il *fallut* que le **président** Franklin D. Roosevelt intervienne pour mettre un terme (temporaire) à ces enquêtes, qui visaient des écrivains tels que Truman Capote ou William Faulkner. (WIKI)
184. En 1980, les États-Unis et 64 autres délégations boycottent les Jeux de Moscou en raison de l'intervention soviétique en Afghanistan. La France ou encore le Royaume-Uni ne se sont pas solidarisés avec ce mouvement et se rendent à Moscou avec quatorze autres nations occidentales. Le Comité olympique américain (USOC) tente de passer outre l'ordre de boycott donné par la Maison Blanche. Il *faut* que le **président** américain Carter menace les athlètes d'interdiction de sortie de territoire pour faire plier l'USOC. (WIKI)
185. Pour ce faire, il part d'un constat : l'apprentissage le plus répandu et qui réussit le mieux, est celui de l'apprentissage de la langue. Or, cet apprentissage se produit naturellement, sans intervention d'enseignant désigné. Il se construit, point barre. Le Logo group va s'inspirer de ce modèle d'apprentissage de la langue (en dehors de l'école) pour tenter de l'appliquer à d'autres domaines. Pour cela, il *faut* que l'**enfant** continue à trouver un sens à ce qu'il fait, que ceci s'accorde avec sa perception et ce qu'il connaît déjà. (WIKI)
186. L'hypothèse de Papert est que la pratique de la programmation structurée aura des conséquences sur le raisonnement de l'enfant face à certaines autres tâches complexes dans la mesure où il y aura transfert de cette manière de décrire et d'aborder un problème, lorsqu'il sera confronté à d'autres situations situées dans un autre champ d'action, au nouveau problème qui se pose à lui. Mais encore *faut-il* pour cela que l'**enfant** accède à cette programmation structurée ! Encore une pierre d'achoppement. (WIKI)

Comme pour les exemples 177 à 182, l'interprétation de *falloir* se lit dans ces occurrences avec la préposition *pour* comme une *nécessité tournée vers un but*.

En somme, lorsque le sujet sémantique est exprimé pour *falloir*, il prend la plupart du temps la forme d'un *animé*. La locutrice exprime une *nécessité*, tournée vers un but, et avec une certaine prise de distance par rapport à une tournure similaire avec *devoir*. En effet, il peut être plus « agressif » de dire *le président doit* plutôt que *il faut que le président*. C'est probablement pour cette raison que la plupart des sujets sémantiques significatifs sont des êtres *animés* pour *falloir* : lorsqu'on exprime sur qui porte cette *nécessité de contrainte*, il vaut mieux mettre de la distance afin de ne pas froisser les personnes ou les instances sur qui portent les *contraintes*. En effet, parmi les stratégies qui permettent de ménager la face de l'interlocutrice, on retrouve chez Brown et Levinson (1987 : 70) l'usage des tournures impersonnelles.

7.1.6. L'AFC et l'analyse arborée appliquée aux sujets nominaux

Maintenant que nous connaissons les profils combinatoires de chacun des verbes modaux avec les sujets nominaux significatifs, nous allons observer de manière plus précise leurs rapprochements et leurs différences à l'aide de l'AFC et de l'analyse arborée. Nous avons déjà pu soulever quelques similarités et divergences à l'aide de l'indice de spécificité mais, comme susmentionné en 6.4.2. et 6.4.3., l'AFC et l'analyse arborée permettent de mettre en exergue en un seul et même graphique les principaux éloignements et proximités entre les items sélectionnés.

Les caractéristiques que nous avons soulevées dans les parties 7.1.3. à 7.1.5. viendront ici étayer les analyses des graphiques. En effet, avant d'entrer dans les résultats, rappelons que la lecture des AFC et de l'analyse arborée est toujours partielle étant donné que les données ne peuvent pas être entièrement représentées dans des graphiques à deux dimensions.

Dans cette section, nous allons nous attarder à nouveau sur les résultats des deux corpus LM et WIKI employés pour illustrer les profils combinatoires de la section précédente. Nous allons dans un premier temps étudier les quatre verbes ensemble sous leurs formes lemmatisées, puis nous attarder sur les résultats qui mettent en parallèle les formes au présent et au conditionnel pour *pouvoir*, *devoir* et *vouloir*. Ces recherches sont d'abord menées sur les AFC puis appuyées par les résultats de l'analyse arborée.

7.1.6.1. Les résultats dans LM pour les formes lemmatisées

La première AFC¹⁶⁵ que nous allons analyser est celle des formes lemmatisées pour les quatre verbes dans le corpus de presse LM.

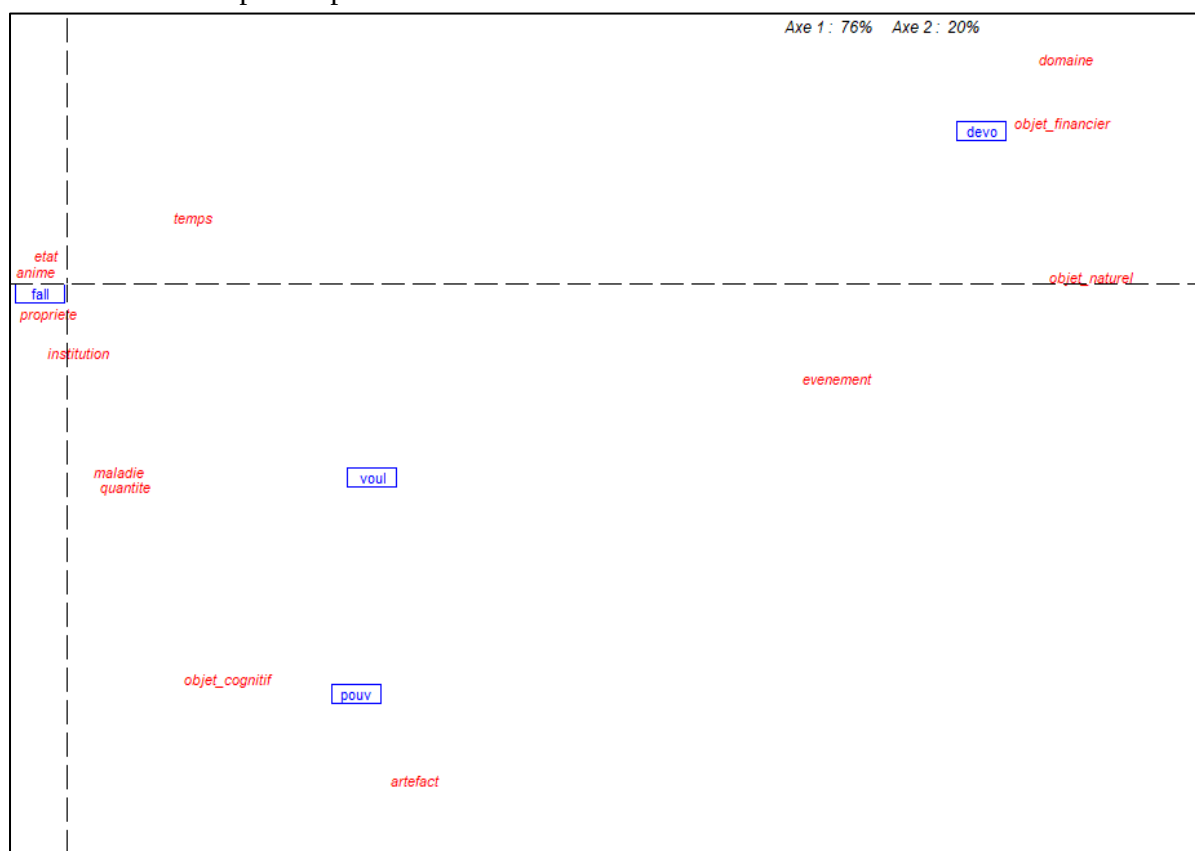


Figure 40 : AFC des modaux sous leurs formes lemmatisées avec les catégories sémantiques des noms en position de sujet grammatical (ou sémantique pour falloir) dans LM.

Que ce soit sur l'axe 1 ou sur l'axe 2, *pouvoir* et *devoir* se retrouvent éloignés l'un de l'autre. Malgré les points communs que nous avons ainsi repérés entre leurs profils combinatoires respectifs sous leurs formes lemmatisées (7.1.3.), ceux-ci présentent finalement plus de divergences que de similarités. Comme nous l'avons soulevé, *pouvoir* se distingue par son attirance pour les sujets de type *objets cognitifs* et *artefacts*. *Devoir*, quant à lui, s'éloigne des autres verbes par ses cooccurents relevant des noms des catégories de *domaines*, d'*objets financiers* et *naturels*. Une telle attraction pour ces catégories ne le distingue pas seulement de *pouvoir* mais également de *vouloir* et *falloir*. En effet, il semble y avoir une dichotomie importante entre d'un côté ces trois verbes et de l'autre *devoir*. Nous pensons que cela pourrait

¹⁶⁵ Celle-ci a été obtenue en construisant le tableau des cooccurrences brutes que chacun des modaux présentait avec les noms en position sujet qui sont ressortis lors du calcul des cooccurrences spécifiques. Ces noms, comme pour les analyses précédentes, ont été regroupés dans les catégories de Salvadori et al. (2021) afin de représenter les rapprochements et éloignements entre les verbes modaux par rapport aux catégories des noms qu'ils attirent en position sujet. Ce même procédé a ensuite été adapté pour les AFC qui suivent.

s'expliquer par les observations que nous avons soulevées tout au long de la partie consacrée aux profils combinatoires de ces verbes : pour *devoir*, nous arrivions toujours à retrouver des effets de sens ou des valeurs modales attestées par les diverses autrices dans les études à son propos – et plus précisément son NS de *nécessité* –, alors que les autres modaux semblent présenter des usages plus distincts de ces modalités traditionnelles, puisque ces effets de sens sont plus souvent inférés que directement déduits du cotexte. Nous avons par exemple pu apercevoir l'expression des *normes* ou des *suppositions* avec *vouloir* lorsque celui-ci est précédé par des *objets cognitifs*. Nous avons aussi illustré une diversité importante d'interprétations sémantiques pour *pouvoir*, ce qui ne le rattache pas à une catégorie modale de prédilection en particulier. Enfin, *falloir* semble surtout atténuer une *contrainte* qui porte sur un *animé* qui pourrait être trop directe si elle était exprimée avec *devoir*. C'est d'ailleurs ces deux verbes qui sont le plus éloignés sur l'axe 1, ce qui ne rejoint pas nos hypothèses : en effet, nous estimions que les deux verbes qui permettent habituellement d'exprimer les valeurs modales déontiques auraient des comportements plus proches que les autres verbes entre eux. En termes de proximité et d'éloignement, nous pouvons aussi procéder à une lecture par cadran (Achat & Marty 2016). Dans cette première AFC, il y a trois regroupements majeurs : le premier en bas à gauche avec *falloir*, le second en bas à droite avec *vouloir* et *pouvoir* et enfin le dernier en haut à droite avec *devoir*. Ces regroupements et éloignements peuvent paraître étonnants car les ressemblances relevées dans la littérature jusqu'à présent concernent soit les verbes pouvant exprimer des modalités épistémiques (donc *pouvoir* et *devoir* principalement) ou alors ceux qui véhiculent des valeurs modales déontiques (en l'occurrence *devoir* et *falloir*).

En observant de plus près le cadran en bas à droite, nous pouvons noter que *vouloir* et *pouvoir* présentent des similitudes notables sur l'axe 1 (qui contient ici 76% des informations du tableau) au vu de leurs positions respectives. Cependant, autour de *vouloir* apparaît un certain « vide » alors que nous pouvons apercevoir les catégories des *objets cognitifs* ainsi que les *artefacts* proches de *pouvoir*. Ces catégories expliquent également le positionnement de *vouloir* à proximité de *pouvoir* mais en même temps proche du centre et donc des sujets *animés*.

Il convient de souligner que les proximités entre les verbes et certaines catégories de noms ne suggèrent pas obligatoirement l'absence d'affinités entre les autres verbes modaux et ces catégories. Au contraire, elle met en évidence, par exemple, la distinction de *pouvoir* par l'intensité particulière de ses associations spécifiques avec certains types de noms. Cette analyse se réalise en considérant l'ensemble des associations avec d'autres catégories, tout en tenant compte également des associations de ces verbes avec ces catégories.

Lors d'une lecture d'une AFC, il peut également être intéressant de se pencher au centre des deux axes. Un point se trouvant à ce croisement représente une non-singularisation par rapport aux paramètres étudiés. Lebart et Salem (1994) parlent d' « un profil moyen » (Lebart & Salem 1994 : 84-85) pour cette position particulière. Ce point ne se démarque ainsi pas par rapport aux paramètres étudiés lorsqu'il est comparé avec d'autres items sur ces mêmes critères. Plus précisément,

les points proches de l'origine de l'axe ne ressortent pas significativement pour observer les contrastes [...] les textes les plus proches de l'origine sont également les plus proches de la norme du corpus. Les autres parties ont donc des particularités plus spécifiques par rapport auxquelles elles se démarquent significativement par rapport aux autres. (Poudat & Landragin 2017 : 108)

Dans notre cas, les *textes* mentionnés par Poudat et Landragin (2017) s'apparentent aux verbes modaux que nous étudions¹⁶⁶.

Dans la figure 42, nous pouvons apercevoir que *falloir* se situe au centre de ces deux axes. Ceci illustre les dires des autrices que nous venons de mentionner : *falloir*, par son manque de cooccurrences avec les sujets nominaux, ne se démarque pas des autres verbes modaux sur ce paramètre. Notons également que les points qui représentent les catégories au centre de l'axe sont *état*, *animés*, *propriété* et *institution*, ce qui signifie que ce ne sont pas sur ces groupes de noms que les verbes se différencient.

¹⁶⁶ Les autrices font ici référence à des études textométriques, c'est pourquoi elles mentionnent les *textes* dans la citation.

Ces observations peuvent être quelque peu nuancées avec l'analyse arborée :

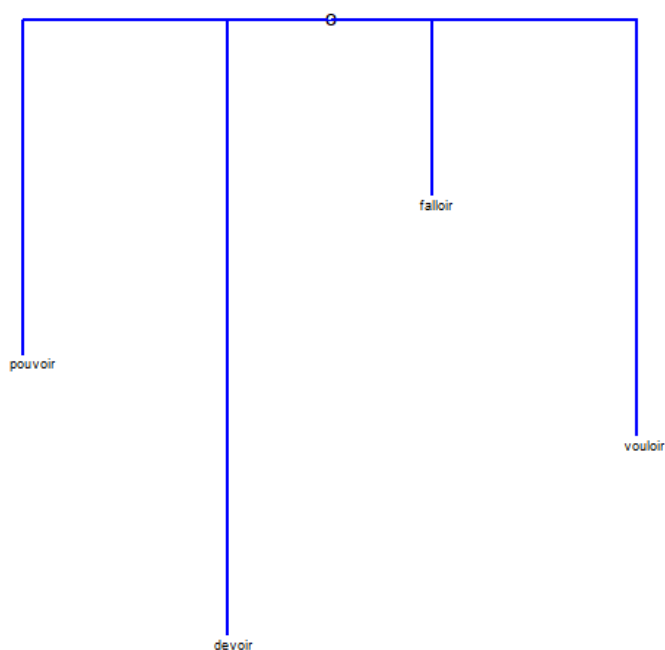


Figure 41 : Analyse arborée des modaux sous leurs formes lemmatisées avec les catégories sémantiques des noms en position de sujet grammatical (ou sémantique pour falloir) dans LM.

Nous pouvons ici rappeler les principes de la lecture d'une analyse arborée qui consiste à observer des regroupements et des oppositions (au même titre que les AFC dans ce sens). Ces regroupements se lisent à travers les nœuds et les oppositions à travers les branches (Mayaffre & Luong 2003 : 6).

Ainsi, le nœud sépare ici *pouvoir* et *devoir* d'un côté et *falloir* et *vouloir* de l'autre. Ces rapprochements ne ressortaient pas à l'aide de l'AFC mais sont finalement assez compréhensibles si on resitue les résultats des profils combinatoires : *pouvoir* et *devoir* attirent tous deux majoritairement des sujets de type *évènement* et *falloir* et *vouloir* des sujets *animés*. Il est aussi important de noter qu'en termes de distance pure, *devoir* et *falloir* sont finalement très proches ici contrairement à ce que nous avons observé sur l'AFC. *Devoir* reste également le verbe le plus éloigné des autres formes dans ce graphique. Nous le remarquons à la taille de sa branche qui est nettement supérieure à celle des autres modaux.

Pour ces premières AFC et analyse arborée, nous pouvons nuancer quelque peu les propos de Mayaffre et al. (2019) cités en 6.4.3. : l'analyse arborée ne « corrobore » ici pas entièrement les résultats de l'AFC : elle les nuance. En effet, il semblerait que les différences que nous avons observées sur l'AFC de la figure 42 sont moins importantes que les points communs qui ressortent avec les profils combinatoires et l'analyse arborée. Ainsi, nous avons soulevé que *pouvoir* et *devoir* se distinguent l'un de l'autre par des attractions avec certaines catégories

sémantiques de noms en particulier – les *artefacts* et les *objets cognitifs* pour *pouvoir*, les noms de *domaine*, *objets financiers* et *naturels* pour *devoir* – mais ils semblent également avoir de nombreux points communs sur la figure 43 puisqu'ils se retrouvent du même côté du nœud. Nous aurions pu nous attendre à plus de différences entre ces deux verbes dont les NS – la *possibilité* et la *nécessité* – représentent les deux extrémités de l'échelle modale au sens traditionnel du terme. Il semblerait que, dans l'usage, cela soit moins fréquemment le cas dans un discours de presse comme l'illustrent ces premiers graphiques forgés au moyen des données du corpus LM. Cette ressemblance pourrait être due à un usage plus fréquent de la valeur modale épistémique dans ce type de texte que nous avons déjà relevé dans les diverses occurrences analysées dans la section 7.1.3. En effet, comme le mentionnent Lochard et Boyer (1998), un des principes de l'écriture de presse « réside dans l'emploi de différentes formes de *modalisations*, qui permettent de prendre ses distances avec le propos tenu par la *source* » (Lochard & Boyer 1998 : 51) et comme nous l'avons soulevé, c'est dans l'expression de leurs valeurs modales épistémiques que ces deux verbes sont employés dans des cotextes similaires. Ces rapprochements et éloignements sont nuancés à l'aide des lectures des AFC et de l'analyse arborée des verbes par temps verbal dans ce même corpus dans la section suivante.

7.1.6.2. Les résultats dans LM pour les formes au présent de l'indicatif et au conditionnel présent

Dans l'AFC suivante, *pouvoir*, *devoir* et *vouloir* sont comparés par rapport à leurs formes au présent de l'indicatif et au conditionnel présent¹⁶⁷. Nous avons omis ici *falloir*, car il possède trop peu de cooccurrences significatives pour les sujets dans la complétive au présent de l'indicatif et au conditionnel présent.

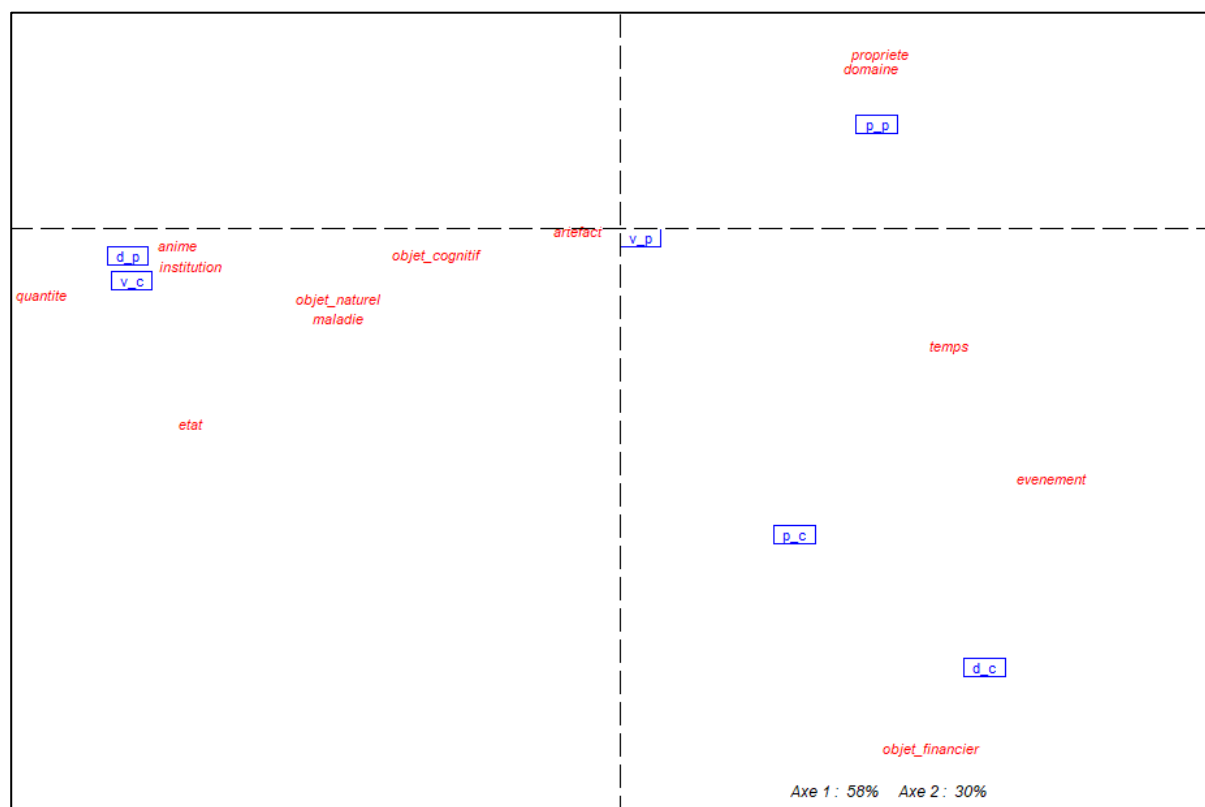


Figure 42 : AFC des verbes *pouvoir*, *devoir* et *vouloir* au présent de l'indicatif et au conditionnel présent avec les catégories sémantiques des noms en position de sujet grammatical dans LM.

L'éloignement le plus important à noter sur l'axe 1 concerne les deux formes de *devoir*. En effet, du côté gauche de l'axe, nous retrouvons *devoir* à l'indicatif présent à proximité de *vouloir* au conditionnel et de l'autre, *devoir* au conditionnel. Sur l'axe 2, c'est cette forme au conditionnel présent de *devoir* qui se trouve la plus opposée à celle de *pouvoir* au présent se situant en haut de l'axe.

Sur l'axe 1, nous remarquons que *pouvoir* au conditionnel présent est proche de sa forme au présent de l'indicatif. Ceci peut montrer une certaine stabilité d'usage du verbe, quelle que soit sa forme. Ce qui ne semble pas être le cas ni pour *devoir*, ni pour *vouloir*, relativement éloignés lorsqu'ils sont conjugués au présent de l'indicatif ou au conditionnel présent. Sur l'axe 2,

¹⁶⁷ *p_p* par exemple pour *pouvoir* au présent de l'indicatif et *p_c* pour *pouvoir* au conditionnel présent, et ainsi de suite.

cependant, les deux formes de *vouloir* sont très proches, ce qui témoigne d'une certaine stabilité de ce verbe également en matière de sujets grammaticaux nominaux attirés.

Au centre des deux axes, nous trouvons cette fois-ci le verbe *vouloir* conjugué au présent de l'indicatif. En comparaison avec les autres formes verbales, il se distingue le moins par rapport aux catégories nominales représentées dans les sujets significatifs de ce corpus. Cependant, nous remarquons sa proximité avec les artefacts et les objets cognitifs, ce qui suggère que ce sont sur ces associations spécifiques que se manifeste une certaine différenciation du verbe *vouloir* par rapport aux autres formes verbales. Cela confirme ainsi nos précédentes observations sur ses utilisations grammaticalisées.

Vouloir au présent se trouve dans le même cadran en bas à droite avec *pouvoir* et *devoir* au conditionnel. Ces deux formes ont un comportement très similaire ce qui peut s'expliquer par leur proximité avec les noms d'évènement que nous avons déjà relevé dans leurs profils combinatoires respectifs (7.1.3.11 et 7.1.3.13). Ceci rejoint également notre hypothèse sur le fait que, lorsque ces verbes expriment la même valeur modale – en l'occurrence ici la valeur modale épistémique – leurs cotextes d'utilisation sont semblables.

Dans le cadran en haut à droite, qui comporte uniquement *pouvoir* au présent, nous pouvons remarquer que ce verbe se rapproche des catégories de noms de *domaine* et de *propriété*. Ce qui différencie ainsi *pouvoir* au présent avec ce type de sujet est la valeur modale descriptive qu'il véhicule dans ces cotextes puisqu'on peut ici comprendre ces phrases comme dénotant des *capacités* de ces *domaines* et *propriétés* :

187. Il avait compris que la **politique peut** devenir, si l'on n'y prend pas garde, une forme du sacré. (LM, 12.04.2010)
188. Seul l'**intérêt peut** pousser la Chine à mettre de côté ses immenses besoins en pétrole iranien pour prendre ses responsabilités de puissance diplomatique montante et se joindre au concert des Nations souhaitant des sanctions. (LM, 17.04.2010)

Cette valeur modale est rarement associée aux autres verbes que nous étudions ici, ce qui pourrait ainsi expliquer la position éloignée de *pouvoir* au présent par rapport aux autres formes du graphique.

Enfin, dans cette AFC nous remarquons également une proximité importante entre *devoir* au présent et *vouloir* au conditionnel dans le cadran en bas à gauche. Ce groupement semble être dû à leur proximité avec les sujets de type *animé* et les *institutions* et donc leurs NS respectifs

qui sont davantage exprimés avec ces catégories-là, NS qui trouvent ainsi des similitudes quant au type de sujet qu'ils attirent.

L'analyse arborée de ces mêmes chiffres corrobore les regroupements et éloignements que nous avons relevés dans l'AFC :

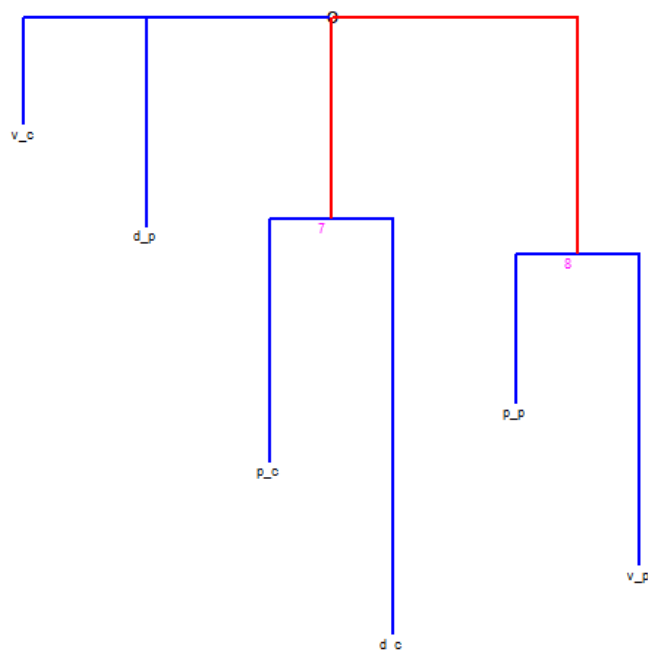


Figure 43 : Analyse arborée des verbes *pouvoir*, *devoir* et *vouloir* au présent de l'indicatif et au conditionnel présent avec les catégories sémantiques des noms en position de sujet grammatical dans LM.

Ainsi, le premier nœud sépare d'un côté *vouloir* au conditionnel avec *devoir* au présent (ce que l'AFC montrait déjà) et de l'autre côté de ce nœud, toutes les autres formes.

Le regroupement entre *pouvoir* et *devoir* au conditionnel se confirme ici. Cependant, concernant *vouloir* au présent, celui-ci se trouve en fin de compte plus proche de *pouvoir* au présent que de *pouvoir* et *devoir* au conditionnel, comme dans l'AFC de la figure 44. C'est peut-être en raison de la proximité très particulière entre *devoir* et *pouvoir* au conditionnel que d'autres verbes sont écartés de ce regroupement. Cette proximité est due aux valeurs modales épistémiques qu'ils semblent exprimer dans ce mode, comme nous l'avons relevé dans les parties consacrées aux cooccurrences.

Notons également que les deux formes de *pouvoir* se trouvent du même côté de la séparation opérée par le nœud principal alors que *devoir* et *vouloir* ont une forme à gauche et à droite de ce nœud. Comme pour l'AFC, cela révèle ainsi une certaine stabilité du comportement de *pouvoir* peu importe le temps verbal auquel il est conjugué – du moins, dans ce corpus avec les sujets sous forme nominale – alors que les effets de sens et valeurs modales divergent davantage pour les autres verbes selon la forme qu'ils revêtent. Les deux formes de *vouloir* sont celles qui

sont les plus éloignées l'une de l'autre, ce qui illustre des comportements différents du verbe selon le temps verbal auquel il est conjugué. Rappelons que nous avons soulevé l'importance des sujets de type *objets cognitifs* au conditionnel et des êtres *animés* au présent, donc, d'un côté une forme plus grammaticalisée du verbe contre, de l'autre, l'expression du NS de manière plus prépondérante.

7.1.6.3. Les résultats dans WIKI pour les formes lemmatisées

En ce qui concerne les résultats pour les formes lemmatisées dans WIKI, l'AFC montre ici d'autres positionnements de ces verbes les uns par rapport aux autres et nous amène ainsi à voir les ressemblances et divergences sous un autre angle :

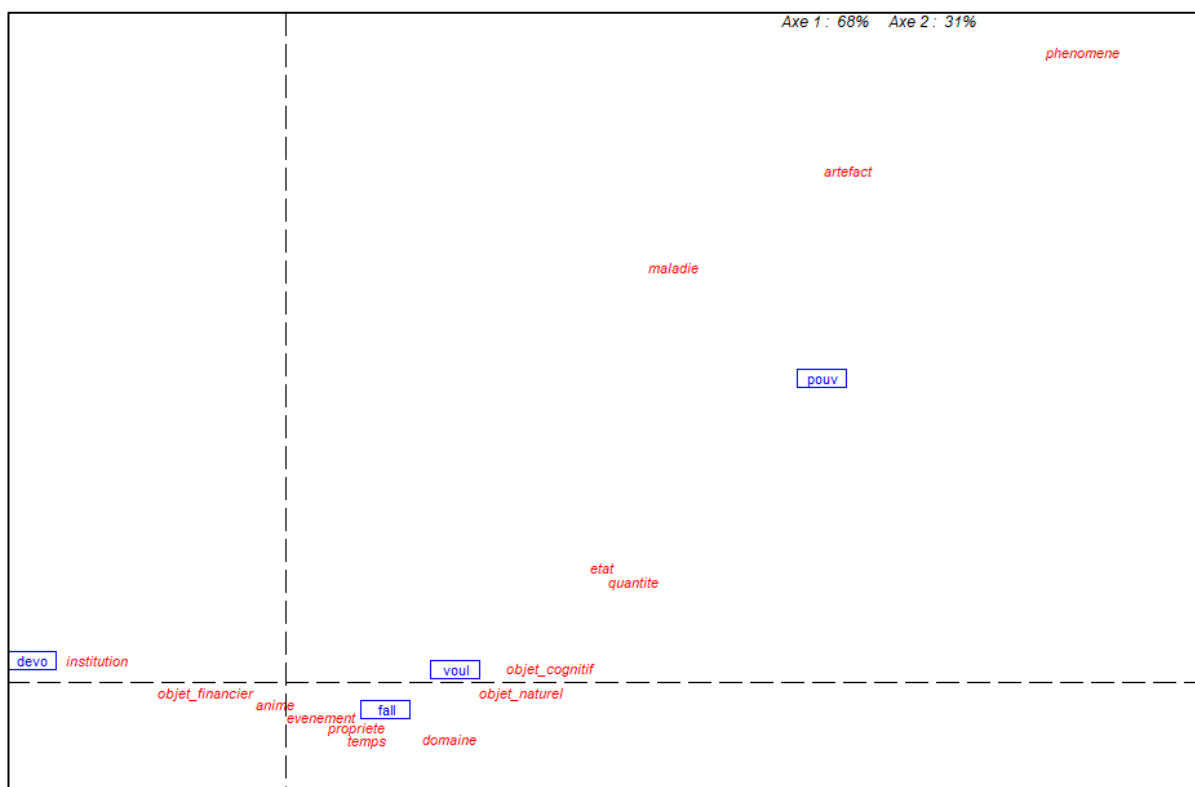


Figure 44 : AFC des modaux sous leurs formes lemmatisées avec les catégories sémantiques des noms en position de sujet grammatical (ou sémantique pour falloir) dans WIKI.

Sur l'axe 1, ce sont les verbes *devoir* et *pouvoir* qui sont les plus éloignés sous leur forme lemmatisée dans WIKI, ce qui peut s'expliquer par leurs profils combinatoires moins proches dans ce corpus que dans LM. L'axe 2 comporte de moins grands extrêmes, car *pouvoir* se trouve en haut de celui-ci, mais en bas, les trois autres verbes se retrouvent relativement au même niveau. Il semble ainsi s'opérer une distance entre *pouvoir* d'un côté et les trois autres modaux de l'autre. *Pouvoir* semble présenter dans ce corpus un comportement qui le démarque véritablement des autres alors que c'était davantage *devoir* qui semblait se singulariser dans le

corpus LM avec ces mêmes paramètres. Pour *pouvoir*, le fait est que, dans ce corpus, nous avons observé qu'il pouvait régulièrement signifier une valeur modale descriptive (7.1.4.3.). Cette valeur modale est rarement exprimée par les autres verbes et c'est ceci qui peut expliquer ici la distinction de *pouvoir* par rapport aux autres verbes.

Les trois modaux sont d'ailleurs proches du centre des deux axes. Le plus centré est *falloir*, ce qui montre qu'il y a une certaine stabilité dans son comportement avec ces paramètres selon le texte dans lequel il est usité. *Devoir* est dans ce corpus plus proche de *falloir* dans son comportement qu'il ne l'était dans LM. Ils se situent toutefois dans deux cadrans distincts, ce qui était déjà le cas dans le corpus de presse. Dans l'environnement de *devoir* se trouvent les catégories des noms de type *institution* et *objet financier*. Le verbe semble donc se différencier des autres par une attraction importante avec des noms du monde politique ou encore économique¹⁶⁸.

Vouloir se situe dans le même cadran que *pouvoir* bien qu'ils soient plus éloignés l'un de l'autre que dans le corpus LM. Ici, *vouloir* est plus proche de *falloir* sur l'axe 1 et également très proche de *devoir* sur l'axe 2. Les catégories nominales qui se trouvent dans son entourage sont les *objets naturels* et *cognitifs* ce qui démontre un nombre important d'usages dans lesquels *vouloir* est grammaticalisé dans ce corpus, ce qui a été corroboré par ses profils combinatoires relevés dans les sections précédentes.

Ces catégories peuvent aussi expliquer pourquoi *vouloir* se trouve dans le même cadran que *pouvoir* : ce verbe semble être attiré par des noms non-animés comme des *objets* – avec la catégorie des *artefacts* – mais aussi des éléments plus abstraits comme les *maladies*.

Concernant l'analyse arborée qui se fonde sur les mêmes chiffres que l'AFC, nous remarquons ici une division similaire à celle que nous avons obtenue pour le corpus de presse :

¹⁶⁸ En effet, parmi les noms d'*institution* et d'*objet financier* ayant un indice de spécificité positif en position de sujet pour *devoir* sous sa forme lemmatisée dans WIKI nous retrouvons *gouvernement, dette, entreprise, pays, somme, tribut, armée, agence, rente, fret, euro, montant*, ou encore *capitale*.

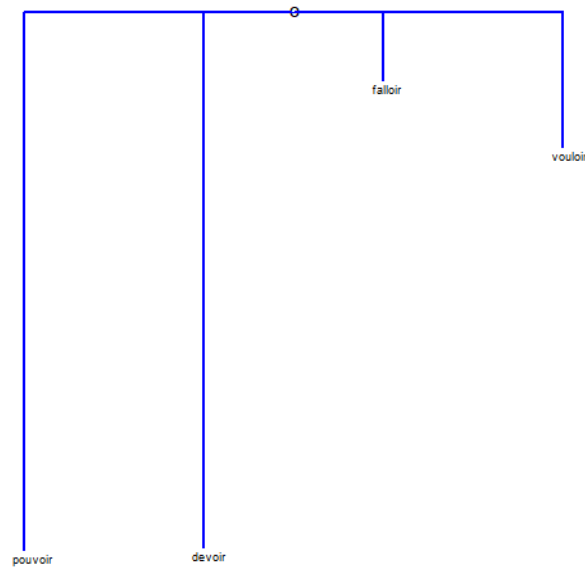


Figure 45 : Analyse arborée des modaux sous leurs formes lemmatisées avec les catégories sémantiques des noms en position de sujet grammatical (ou sémantique pour falloir) dans WIKI.

Nous retrouvons ainsi d'un côté du nœud *pouvoir* et *devoir* et de l'autre *falloir* et *vouloir*. Comme pour LM, cette analyse arborée ne corrobore pas toutes les observations relevées pour l'AFC. Dans la figure 47, *pouvoir* et *devoir* semblent avoir des points communs – vu qu'ils se retrouvent tous les deux du même côté de la division effectuée par le nœud – mais la distance entre leurs feuilles est relativement importante – au vu de la taille respective de leurs branches. *Falloir* et *vouloir* se retrouvent de l'autre côté du nœud et plus proches de celui-ci que *pouvoir* et *devoir*. Ce type de position peut être comparée à celle dans l'AFC : leurs comportements sont moins distinctifs de la norme que ceux de *pouvoir* et *devoir*. En ceci, nous voyons donc bien que *pouvoir* se distingue ainsi nettement des autres verbes en étant celui qui est le plus éloigné de ce nœud ce qui confirme ici notre lecture de l'AFC dans ce même corpus.

7.1.6.4. Les résultats dans WIKI pour les formes au présent de l'indicatif et au conditionnel présent

Les résultats pour les formes conjuguées, soit le présent de l'indicatif et le conditionnel présent, nous permettent d'étayer les ressemblances et les divergences entre les verbes dans ce corpus :

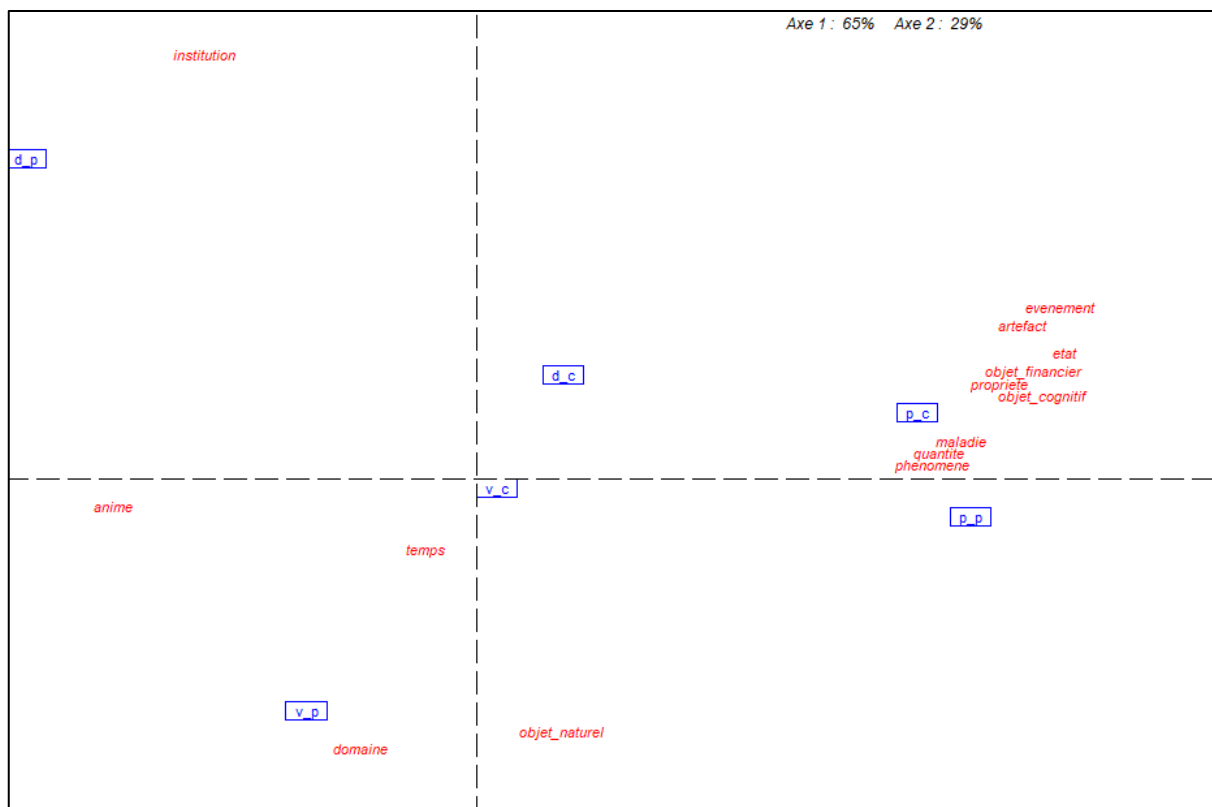


Figure 46 : AFC des verbes pouvoir, devoir et vouloir au présent de l'indicatif et au conditionnel présent avec les catégories sémantiques des noms en position de sujet grammatical dans WIKI.

Les deux formes de *pouvoir* se trouvent du côté droit de l'axe 1 et s'opposent ainsi à *devoir* au présent de l'indicatif. D'ailleurs, cette forme est la plus éloignée de toutes sur l'axe 2, et est opposée à *vouloir* à l'indicatif présent qui se trouve en bas de cet axe. *Devoir* au présent de l'indicatif se trouve à proximité de noms de type *institution* et, un peu plus éloignés, des *animés*. Comme nous l'avons relevé (7.1.4.5.), avec ce type de sujet, *devoir* exprime, dans la grande majorité des cas, la valeur modale déontique. Ceci pourrait ainsi expliquer son éloignement drastique des autres formes représentées dans ce graphique : en effet, nous avons soulevé des valeurs modales et effets de sens moins traditionnels pour *pouvoir* et *vouloir* qui les éloignaient ainsi de leurs NS respectifs.

Au centre du graphique se situe *vouloir* au conditionnel présent. Rappelons que dans LM il s'agissait également de ce verbe mais au présent de l'indicatif. *Vouloir* au conditionnel se trouve très proche sur l'axe 1 de sa forme au présent de l'indicatif – en bas à gauche – mais encore plus proche de *devoir* au conditionnel présent. Celui-ci, comme pour le corpus de presse, se situe dans le même cadran que *pouvoir* au conditionnel présent : leur proximité peut, une nouvelle fois, s'expliquer par leurs cotextes similaires qui leur confèrent généralement une valeur modale épistémique.

Au conditionnel présent, *pouvoir* conserve une proximité notable avec sa forme au présent de l'indicatif. Les deux sont, d'ailleurs, associés à un large éventail de catégories sémantiques de noms en position sujet. À la lumière des observations réalisées jusqu'à présent, nous considérons que cette attraction envers de nombreuses catégories peut être expliquée par la polysémie de ce modal. En effet, comme déduit à partir de nombreux exemples dans les sections 7.1.3. et 7.1.4., les interprétations de *pouvoir* sont toujours très diversifiées, indépendamment du type de sujet significativement attiré par le modal. Étant donné son absence de préférence marquée pour un effet de sens particulier en fonction du type de sujet, *pouvoir* a la capacité d'attirer une diversité importante de catégories sémantiques dans cette position.

Dans l'analyse arborée de ces mêmes chiffres, la plus grande différence semble pourtant se situer ailleurs :

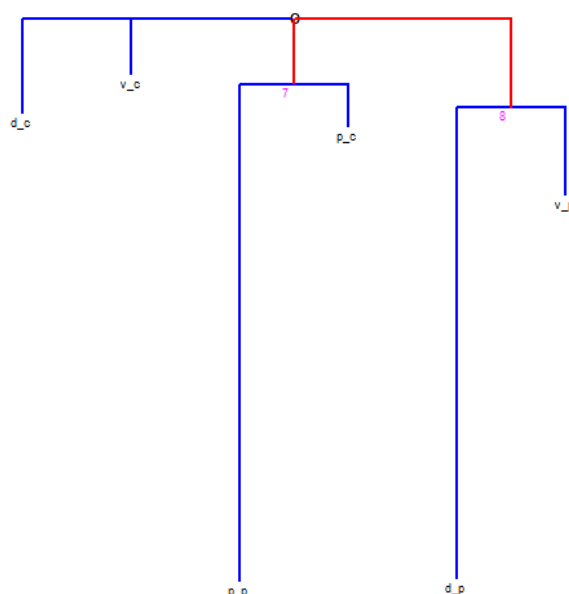


Figure 47 : Analyse arborée des verbes *pouvoir*, *devoir* et *vouloir* au présent de l'indicatif et au conditionnel présent avec les catégories sémantiques des noms en position de sujet grammatical dans WIKI.

Une dichotomie s'opère dans ce graphique avec, de manière isolée du côté gauche, *devoir* et *vouloir* au conditionnel. Cette séparation peut s'expliquer par leur proximité avec le centre dans l'AFC en figure 48, et donc leur banalité, qui les isole d'un côté du graphique. De l'autre côté de ce nœud principal, nous retrouvons les regroupements déjà repérés sur l'AFC : ainsi, les deux formes de *pouvoir* sont réunies dans un même embranchement, et *devoir* et *vouloir* au présent de l'indicatif dans un autre. Il semblerait ainsi que pour ces deux verbes, les temps et modes verbaux jouent davantage un rôle dans le type de sujets qu'ils attirent que pour *pouvoir* qui paraît agir de la même manière qu'elle que soit la forme qu'il revêt.

Synthèse sur les associations entre les verbes modaux et les sujets nominaux

Ce premier pan de l'analyse avait pour but d'observer comment les sujets grammaticaux sous forme de noms pouvaient influencer les propriétés sémantiques des verbes modaux. Nous avons commencé par rappeler que, même si *pouvoir* et *devoir* étaient traditionnellement considérés comme des verbes à *héritage* et *vouloir* comme un verbe à *partage*, cette dichotomie n'avait pas lieu d'être au vu des nombreux emplois grammaticalisés conférés à *vouloir* lorsqu'il est précédé par un sujet non-animé. Nous avons aussi rappelé que dans plusieurs études, l'influence du sujet sur l'interprétation sémantique des verbes modaux avait été mentionnée. Cette contradiction a été donc le point de départ de notre analyse : est-ce que les sujets ont une influence sur l'interprétation du modal, et si oui, laquelle ? La dichotomie entre verbes à *héritage* et à *partage* a-t-elle encore lieu d'être ?

Pour classer nos résultats, nous nous sommes appuyée sur la typologie des noms présentée par Salvadori et al. (2021). Cette typologie a fourni une base pour catégoriser les noms sujets en fonction de leurs caractéristiques sémantiques, offrant une structure claire pour l'analyse des préférences combinatoires des verbes modaux.

A partir de ces différentes réflexions théoriques, nos analyses ont révélé des différences et ressemblances significatives dans les comportements combinatoires des verbes modaux au sein des corpus étudiés. Ainsi, dans le corpus de presse, *pouvoir* et *devoir* ont des comportements très similaires et ils attirent en particulier des noms que nous avons catégorisés comme étant des *événements*. Cette préférence était plus ou moins marquée selon la forme analysée, avec davantage de similarités pour les profils combinatoires du conditionnel présent que pour le présent de l'indicatif.

Nous avons aussi pu soulever, en observant les contextes d'utilisation, que les effets de sens et valeurs modales étaient nombreux pour ces verbes, en particulier pour *pouvoir* dont la lecture sémantique est variée, quel que soit le type de sujet observé. *Devoir*, lui, a un comportement plus stable : la valeur modale épistémique apparaît surtout dans la presse et au conditionnel, mais dans le reste des occurrences, c'est majoritairement la valeur modale déontique qui ressort sous diverses formes.

Vouloir, lui, a également présenté des particularités par rapport aux caractéristiques qui lui sont habituellement attribuées dans les analyses à son propos. Certes, bien souvent le NS d'*intention/volonté* ressort dans ses usages – surtout avec des noms *animés* – mais les résultats ont révélé que sa grammaticalisation est également très importante. Ainsi, il apparaît très souvent avec des noms classés parmi les *objets cognitifs* dans la classification adoptée dans ce

travail et ces associations-là amènent une lecture différente de *vouloir*. Dans ce type d'usage, il permet à la locutrice d'émettre une certaine distance sur des *normes* ou des *suppositions*, interprétation que nous avons pu étayer à l'aide des théories pragmatiques sur l'inférence que nous avons présentées dans la section 4.1. du travail mais également en nous fondant sur l'ouvrage de Brown et Levinson (1987) sur la politesse.

Nous avons également observé quelques variations de la lecture sémantique selon le genre textuel dans lequel ces verbes modaux sont utilisés. *Vouloir* est plus souvent grammaticalisé dans le corpus encyclopédique, tout comme *pouvoir*, qui s'éloigne de son NS de *possibilité* pour exprimer plus fréquemment des valeurs modales descriptives, notamment lorsqu'il est associé à des *objets* en position de sujet.

Il serait ainsi très difficile d'annoter automatiquement les occurrences de tel verbe avec tel sujet en leur attribuant une certaine valeur modale ou un effet de sens spécifique au vu des grandes divergences à l'intérieur même d'un paradigme tel que *NOM D'ÉVÉNEMENT + DEVOIR* par exemple. Ces premières observations nous amènent à rappeler quelques généralités sur l'annotation sémantique, soit qu'elle est complexe et difficile à automatiser :

Malgré l'importance et l'utilité de l'annotation sémantique pour de nombreux domaines de la recherche en linguistique (de corpus) – traduction automatique, recherche d'information, analyse de contenu, traitement de la parole, recherche discursivo-pragmatique sur l'ironie, approches basées sur les corpus en lexicographie, etc. – il faut garder à l'esprit que l'annotation sémantique est une tâche extrêmement chronophage et gourmande en ressources. Alors que les humains semblent éprouver très peu de difficultés à accéder et à comprendre le sens approprié d'un mot dans des contextes communicatifs naturels, suffisamment bien pour que la communication ne soit pas compromise – que ce soit littéral ou métaphorique/idiomatique – les humains chargés d'annoter les sens des mots en contexte sont moins souvent d'accord entre eux que ce à quoi on pourrait s'attendre [...], comme quiconque ayant déjà tenté d'annoter les sens d'un mot pourra le confirmer. (Gries & Berez 2017 : 386)

Ceci est d'autant plus difficile dans la distinction des valeurs modales ou des effets de sens des marqueurs modaux où une *probabilité* et une *possibilité* peuvent être très proches dans leurs usages et ainsi être distinguées ou non selon les annotatrices¹⁶⁹. En outre, ce sont également des procédés pragmatiques qui entrent en jeu dans l'interprétation des modaux, et non pas seulement le cotexte dans le sens des mots qui précèdent et suivent le modal, ce qui accentue

¹⁶⁹ Cette difficulté d'attribuer des valeurs modales aux verbes modaux a déjà été relevée notamment par Montrichard et al. (2022) car les contours de ces catégories modales sont flous et qu'en plus les marqueurs modaux sont polysémiques.

cette difficulté car ces procédés ne sont pas repérables de manière automatique avec les outils de la linguistique de corpus.

Les AFC et l'analyse arborée ont permis de mettre sur un même plan toutes ces remarques afin de relever les ressemblances et les divergences des verbes modaux sur la base des sujets nominaux qu'ils attirent. *Pouvoir* et *devoir* se distinguent l'un de l'autre en s'associant à des catégories sémantiques éloignées, même si les points communs – notamment leurs valeurs modales épistémiques – les rapprochent aussi significativement. *Devoir* et *vouloir* semblent d'ailleurs plus proches, surtout au présent, que les autres verbes étudiés ici. *Devoir* s'éloigne aussi significativement de *falloir* même si ce sont deux verbes prototypes de la valeur modale déontique.

Enfin, tout au long de cette analyse, nous avons soulevé l'idée que le sujet nominal ne joue pas un rôle aussi important que prétendu dans la lecture sémantique de ces verbes. C'est plutôt le temps ou le mode verbal auquel les modaux sont conjugués, ainsi que d'autres éléments contextuels, qui permettent d'inférer une interprétation plutôt qu'une autre. Nous allons examiner dans la section suivante si la complémentation sous forme de verbe à l'infinitif a davantage d'influence sur l'interprétation sémantique de ces verbes.

7.2. La complémentation verbale sous forme de verbe à l'infinitif

Par rapport aux sujets grammaticaux – qui sont propres à tous les verbes – la complémentation verbale sous forme de verbe à l'infinitif singularise les modaux par rapport à d'autres types de verbes. Abeillé et al. (2021b) expliquent que les verbes modaux, tout comme les auxiliaires, « [...] ne peuvent à eux seuls décrire un type de situation et doivent se combiner avec une autre forme verbale » (Abeillé et al. 2021b : 131). Ces complémentations verbales permettent de distinguer les effets de sens ou les valeurs modales des verbes. En effet, « [...] tous les verbes ne sont pas susceptibles d'être employés avec n'importe quel verbe modal. Les différents verbes modaux exercent une certaine sélection vis-à-vis de leur verbe constructeur » (Chu 2008 : 69). Cependant, malgré cette importance du verbe à l'infinitif qui suit le modal, les études sur ces associations sont moindres¹⁷⁰.

Dans cette partie, comme pour la précédente consacrée aux sujets de nature nominale, nous allons vérifier si ces remarques concernant les interprétations sémantiques peuvent être

¹⁷⁰ Pour l'influence du verbe à l'infinitif en position de complémentation verbale, voir 4.2.4. pour *pouvoir*, 4.3.7. pour *devoir* et dans une moindre mesure les remarques sur la grammaticalisation de *falloir* (4.4.5.) et de *vouloir* (4.5.6.).

vérifiées dans nos corpus, mais également observer si des catégories particulières de verbes s’associent davantage à un modal plutôt qu’à un autre, et donc ce que ces cooccurrences spécifiques peuvent ainsi révéler des propriétés sémantiques des modaux étudiés ici.

Concernant nos hypothèses de départ, nous pensons que les ressemblances entre *devoir* et *falloir* peuvent ici être importantes car ils représentent tous les deux des verbes prototypiques de la valeur modale déontique et qu’ils modalisent ainsi potentiellement le même type de verbe. Les profils combinatoires de *pouvoir* et *devoir* pourraient à nouveau se ressembler lorsque ces verbes sont au conditionnel présent et qu’ils expriment majoritairement une valeur modale épistémique. En effet, vu que leurs cotextes d’utilisation étaient relativement similaires lorsque nous nous sommes penchée sur l’étude des sujets, cette ressemblance pourrait à nouveau émerger ici. Le fait que *falloir* et *vouloir* n’aient pas nécessairement besoin d’une complémentation verbale pour être énoncé pourrait également les rendre proches. Étant donné que les remarques concernant ces associations sont finalement peu fréquentes dans les études à leur sujet, nous n’avons pas d’autres hypothèses particulières à tester ici et proposons surtout une approche *corpus driven*.

Comme pour les sujets nominaux, nous allons analyser les résultats de cette section en faisant appel à un regroupement sémantique des verbes, celui proposé par Dubois et Dubois-Charlier (1997) que nous présentons dans la section suivante.

7.2.1. Le dictionnaire *Les Verbes Français* de Dubois et Dubois-Charlier (1997)

La sémantique des verbes est une thématique qui semble *a priori* plus complexe que la sémantique des noms. Cette complexité pourrait être due à leurs caractéristiques syntaxiques particulières :

Le verbe [...] a longtemps été considéré comme un élément syntaxique, ne renvoyant pas, par lui seul, à un concept. Il n’y a pas de correspondance directe lexème “verbe” - réalité du monde, autrement dit pas de correspondance “signifiant” - “signifié”. N’étant pas un élément conceptuel, mais syntaxique, les verbes ont été “abandonnés” aux linguistes ainsi qu’aux spécialistes de la syntaxe et ont été considérés comme des éléments nécessitant d’être accompagnés par des noms “pour signifier”. Un verbe se doit donc d’être inséré dans une chaîne de mots, une phrase, pour être un vecteur de signification. Plus précisément, il doit être accompagné de ses rôles thématiques. C’est le plus souvent la chaîne “agent – verbe – patient” qui permet de renvoyer à une réalité. Dans cette conception, le verbe est placé au centre de la phrase, convoquant les rôles thématiques autour de lui. La signification, dépendante de l’assemblage du verbe et de ses rôles thématiques, est construite grâce à l’organisation syntaxique de ces éléments. Dans ce cadre, il est donc impossible de parler de

la sémantique du verbe. Ainsi, le verbe ne renvoyant pas à des représentations cognitives, laisse la fenêtre de l'esprit fermé. (Pariollaud 2008 : 13)

Malgré cette difficulté de représenter sémantiquement un verbe, les classifications de ce type sont nombreuses. Les personnes ayant travaillé sur les catégories sémantiques des verbes les regroupent selon

[...] un type de situation, c'est-à-dire un état, une activité ou un événement. Les verbes qui décrivent un état sont par exemple des verbes de sentiment (*aimer*), de possession (*posséder*) ou d'état mental (*connaître*). Les verbes qui décrivent un événement sont notamment des verbes de déplacement (*courir*), des verbes de communication (*parler*) ou d'autres verbes d'action (*frapper*, *manger*). (Abeillé et al. 2021b : 130)

Abeillé et Koenig (2021b) proposent ainsi trois grandes classes sémantiques des verbes qui sont les verbes d'état, d'activité et d'évènement qui peuvent être divisibles en plusieurs sous-classes, pour un total de 20 catégories¹⁷¹. Ces classes présentent l'avantage d'offrir une granularité variable, permettant de classer les verbes de manière détaillée ou plus générale en fonction des critères sémantiques choisis.

Nous pouvons aussi citer la classification de Desclés (1999) qui se fonde sur des propriétés aspecto-temporelles des verbes afin d'établir ses classes sémantiques. Il présente ainsi trois grandes classes qui sont les schèmes *statiques*, *processuels* et *événementiels*¹⁷². Cette classification s'éloigne d'une vision sémantique des verbes et est davantage orientée vers les réalisations ou non des actions dénotées par ceux-ci. Ceci n'est pas sans rappeler la classification de Vendler (1957) qui considère comment chaque type de verbe se rapporte au temps¹⁷³.

¹⁷¹ Ces catégories sont les suivantes : pour les verbes d'état nous avons les verbes d'identité, de localisation, de mesure, modaux, de possession ou d'inclusion et de sentiment. Pour les verbes d'activité nous avons les verbes d'activité intellectuelle, d'activité physique, de mouvement sur place et de perception. Enfin, pour les verbes d'évènement nous pouvons retrouver les verbes aspectuels, de communication, de contact, de déplacement, d'émission, instrumentaux, météorologiques, d'ordre et d'influence, de transfert de possession et des verbes de changement d'état physique.

¹⁷² Les schèmes statiques sont des schèmes descriptifs comme des verbes d'identification, de localisation spatiale ou temporelle, d'attribution de propriété ou d'état mental/physique. Il y a également des verbes d'activité de type être en train de faire quelque chose ou encore ce qu'il appelle des schèmes cinématiques virtuels, comme les verbes se trouver, se tenir, etc. Pour les schèmes processuels il y a des processuels métaboliques, comme les verbes conservateur d'énergie (*rouler*, *croire*, *aimer*), des verbes orientés ou non vers une action particulière (*courir vers* vs. *courir*), des énergies internes comme *grandir* ou encore des processuels qui ne conservent pas l'action décrite comme *glisser* (la voiture ne glisse pas éternellement par exemple). Enfin, les verbes respectant un schème événementiel sont des verbes de mouvement comme *entrer*, des verbes d'achèvement comme *boire*, transformatifs d'objets comme *apparaître* ou encore *écrire* ou *détruire*.

¹⁷³ Vendler (1957) classe ainsi les verbes en quatre catégories : les verbes d'activité, d'accomplissement, d'état ou encore de réalisation. Chaque verbe est ainsi classé par son rapport au temps : soit en continu, soit en atteignant un point terminal, soit en persistant, soit en se produisant à un moment précis.

Fløttum et al. (2007) proposent une classification sommaire des verbes dans leur ouvrage sur le pronom *on*. Elles expliquent que, pour les verbes,

les classifications lexico-sémantiques ou sémantico-pragmatiques ne sont pas les plus évidentes et faciles. Sans entrer dans le détail de la justification de la classification proposée, nous précisons simplement qu'elle est faite afin d'identifier quelques grandes classes de verbes se combinant avec *on* dans l'article de recherche. Un problème supplémentaire est lié au fait que beaucoup des verbes sont polysémiques et peuvent ainsi avoir des sens différents selon les contextes où ils apparaissent. (Fløttum et al. 2007 : 103)

C'est donc uniquement dans le but de classer les valeurs du pronom *on* qu'elles optent pour sept classes distinctes¹⁷⁴. Il convient de noter que les chercheuses soulignent la complexité de la polysémie des verbes, qui peut poser un défi lors de leur classification dans des catégories sémantiques spécifiques. Cependant, elles ne fournissent pas de détails sur la méthodologie utilisée pour surmonter cette difficulté. En outre, il est pertinent de mentionner que la plupart de leurs propositions de catégories sémantiques sont centrées sur les verbes de *communication*, étant donné leur fréquente association avec le pronom *on*.

Mathieu-Colas (2006) propose une classification fondée à la fois sur des critères sémantiques et sur des critères syntaxiques. Ainsi, « [c]haque classe offre une série de propriétés communes, par exemple, pour les verbes de “transfert d'objet”, un schéma à trois arguments » (Mathieu-Colas 2006 : 4). Les classes verbales de Mathieu-Colas (2006) sont au nombre de 14¹⁷⁵.

L'intérêt de classer les verbes permet « au lieu de décrire, au coup par coup, verbe après verbe, la multitude des emplois, [...] d'étudier de façon systématique des ensembles sémantiques homogènes, pour mieux en analyser les propriétés » (Mathieu-Colas 2006 : 5). C'est ce principe-là que nous désirons appliquer à notre analyse. En effet, ces classes nous permettent de cerner de façon plus précise mais aussi plus succincte les propriétés sémantiques de *pouvoir*, *devoir*, *falloir* et *vouloir*.

C'est aussi en faisant des liens entre syntaxe et sémantique que le dictionnaire *Les Verbes français* (désormais *LVF*) de Dubois et Dubois-Charlier (1997) a été construit. Il « repose sur l'hypothèse qu'il y a adéquation entre les schèmes syntaxiques de la langue et l'interprétation sémantique qu'en font les locuteurs » (François et al. 2007 : 9). Cette classification semble être

¹⁷⁴ Ces sept classes sont les verbes *de recherche, discursifs, cognitifs, de prise de position, de locution, de perception* ou encore la catégorie qu'elles nomment *autres* qui est un « groupe fourre-tout où sont réunis verbes de volonté, de sentiment et autres » (Fløttum et al. 2007 : 104).

¹⁷⁵ On y trouve ainsi des verbes *d'états élémentaires et de modalités, de cognition, de langage et communication, de relations sociales, d'actions et d'activités, psychologiques, relatifs au corps, de rapports à l'espace, d'événements et de phénomènes, de changements d'état, d'apparition/disparition, d'unité et de diversité, d'« ayance »* (ce qui peut se rapprocher des verbes de possession) et de *temps*.

la plus complète à ce jour et surtout la plus fonctionnelle pour une classification semi-automatique des verbes du français. Toutes les classifications que nous avons présentées ci-dessus ont leur propre intérêt mais le *LVF* est le seul ouvrage à disposer d'une version électronique¹⁷⁶ et d'une version téléchargeable en format XML¹⁷⁷. Ceci nous permet d'accéder facilement et automatiquement aux diverses données qu'il comporte, et ce, pour tous les verbes qui ont été classés par les autrices¹⁷⁸. Pour toutes les autres classifications présentées ici, nous aurions dû déterminer par nous-mêmes à quelles classes appartiennent les verbes qui apparaissent dans nos résultats, comme c'était le cas pour les noms qui se trouvaient en position sujet dans la section précédente avec la classification des noms Salvadori et al. (2021).

Le *LVF* comporte les 14 classes sémantiques suivantes :

C	communication,	N	munir, démunir
D	don, privation	P	verbes psychologiques
E	entrée, sortie	R	réalisation, mise en état
F	frapper, toucher	S	saisir, serrer, posséder
H	états physiques et comportements	T	transformation, changement
L	locatif	U	union, réunion
M	mouvement sur place	X	verbes auxiliaires

Figure 48 : Les 14 classes de Dubois et Dubois-Charlier (1997) présentées dans François et al. (2007).

Le fonctionnement de ce dictionnaire est en somme assez simple. Ainsi,

Le dictionnaire contient pour chaque verbe des informations intéressantes, entre autres, le sens (synonyme, définition), le domaine d'emploi et le niveau de langue, la syntaxe (régime, nature des sujets et des compléments), et des phrases simples illustrant le sens et la construction. (Vivès 1998 : 66)

Pour le verbe *habiter* par exemple, nous avons les quatre entrées suivantes :

¹⁷⁶ A consulter sur la page suivante : http://rali.iro.umontreal.ca/LVF_DEM/LVF-DEM-req.html

¹⁷⁷ A télécharger sur la page suivante : http://rali.iro.umontreal.ca/LVF_DEM/LVF-XML.html#telechargement

¹⁷⁸ Ce qui correspond à plus de 25'000 entrées pour 12'000 verbes.

Mot^	no	gr	lex	Domaine	Classe	Opérateur	Const.	Sens	Phrases	Adjectifs	Noms
habiter	1	1aZ	1	LOC	L1a.1	(#) [lc.qp] dans logement	A11 T1300	demeurer quelque part	On h~ en banlieue. On h~ une maison en banlieue, la banlieue.	habitable inhabitable habité inhabité habitant	habitation
habiter	2	1aZ	5	ZOO	L1a.1	(animal) [lc.qp] gîte/nid	A21	vivre	Les oiseaux h~ dans les arbres.	habité inhabité habitant	habitation
habiter	3	1aZ	5	LOC	S3g.1	(#) [grp] lieu	T1300	hanter	Ces êtres étranges h~ cette planète. La maison est h~.	habité inhabité habitant	
habiter	4	1aZ	5	PSYt	S4e.2	(abstrait) [grp.mens]	T3906	hanter, être présent quelque part	La vengeance h~ son coeur. La paix n'h~ plus le monde.		

Figure 49 : Entrées pour le verbe habiter dans le LVF électronique.

Les colonnes qui vont nous intéresser pour la classification automatique des résultats sont celles intitulées *no*, *lex*, et *classe*. Les colonnes *domaine*, *sens* et *phrases* vont nous aider à déterminer le sens exact du verbe en cotexte en cas de doute.

Le *no* et le *lex* sont étroitement liés. Ainsi, *no* correspond, à peu de choses près, au numéro d'entrée dans le dictionnaire, le 1 étant le plus évident. Le *lex*, est en quelque sorte lié à la fréquence. Ainsi, nous pouvons lire sur la page « Aide » du LVF¹⁷⁹ que le numéro 1 correspond au langage « courant » et que le numéro 6 est issu du langage « spécialisé »¹⁸⁰.

Nous aurions pu classer tous les verbes que nous trouvions en cooccurrence spécifique de manière manuelle mais cela aurait été chronophage, comme ça a été le cas pour les noms en position sujet. Afin d'automatiser notre classification, nous avons téléchargé une version .csv du dictionnaire et attribué à nos résultats la catégorie avec le numéro 2 dans la colonne *no*. Ainsi, pour *habiter*, nous l'avons classé comme un verbe *locatif* car la lettre L apparaît dans la colonne *classe* attribué au 2 de la colonne *no*. Ce choix a été fait car, trop souvent, le sens en *no* 1 était trop littéral par rapport à l'emploi dans nos corpus journalistiques et encyclopédiques. Les effets de sens des chiffres au-delà du numéro 2 étaient parfois trop peu fréquents dans nos corpus.

C'est donc une fois ce tri des résultats effectués que nous avons procédé à une analyse que nous présentons dans les pages qui suivent.

7.2.2. Préambule à la recherche

Avant de présenter les résultats, nous établissons ici les fondements nécessaires pour une compréhension approfondie des graphiques et chiffres exposés dans les sections à venir.

¹⁷⁹ La page est à consulter sur le lien suivant : http://rali.iro.umontreal.ca/LVF_DEM/Aide/index.html#LVF-interrogation

¹⁸⁰ Nous remercions Guy Lapalme qui a répondu à nos questions afin de pouvoir préciser ici les diverses significations de ces colonnes.

Tout d'abord, contrairement aux noms, nous n'avons ici qu'un seul grain d'analyse, soit les 14 catégories de Dubois et Dubois-Charlier (1997) présentées dans la figure 49¹⁸¹. Cette classification nous permet d'explorer les regroupements sémantiques des verbes émergeant dans les cooccurrences significatives avec nos modaux. Nous examinerons ensuite de manière détaillée certaines de ces catégories en analysant les occurrences dans le corpus.

Parmi ces catégories, nous incluons celle des *auxiliaires*, qui constitue un ensemble hétérogène. En plus des verbes *être* et *avoir*, elle englobe des verbes impersonnels, inchoatifs, résultatifs et modaux tels que *devenir*, *cesser*, *laisser*, *tarder*, *achever*, etc. Nous avons décidé de conserver cette catégorie dans notre analyse, mais en excluant les auxiliaires de base, *être* et *avoir*, en raison de leur faible pertinence sémantique pour notre étude sur le comportement des verbes modaux. Leur fréquence élevée avait tendance à survaloriser cette catégorie au détriment d'autres groupes de verbes plus riches sémantiquement¹⁸².

Contrairement à notre analyse des noms en position sujet, nous prenons en compte les résultats des quatre modaux étudiés, étant donné que la nature impersonnelle de *falloir* n'influe pas sur cette composante, contrairement à ce qui a été observé dans l'étude des sujets.

7.2.3. Les résultats dans LM

Nous allons commencer par observer les résultats dans le corpus journalistique que nous investiguons pour cette recherche, le corpus LM.

7.2.3.1. Les répartitions des verbes en cooccurrences significatives pour les formes lemmatisées

En tant que lemmes dans le corpus LM, les modaux attirent tous prioritairement des verbes de *communication* et des verbes *psychologiques*.

¹⁸¹ Dubois et Dubois-Charlier (1997) proposent aussi des catégories plus fines dont nous n'avons pas pris compte dans cette étude, car nous nous retrouverions avec 54 classes en prenant en compte le deuxième niveau de classification, 246 avec le troisième et 533 avec le quatrième, ce qui nous semblait être relativement trop détaillé pour une analyse de notre envergure.

¹⁸² Nous avons conscience que ce choix omet ainsi les usages que nous avons qualifiés de tournures passives dans notre analyse des sujets grammaticaux. Il serait pertinent dans un travail ultérieur d'inclure ces cooccurrences et de pouvoir les étudier de manière plus détaillée.

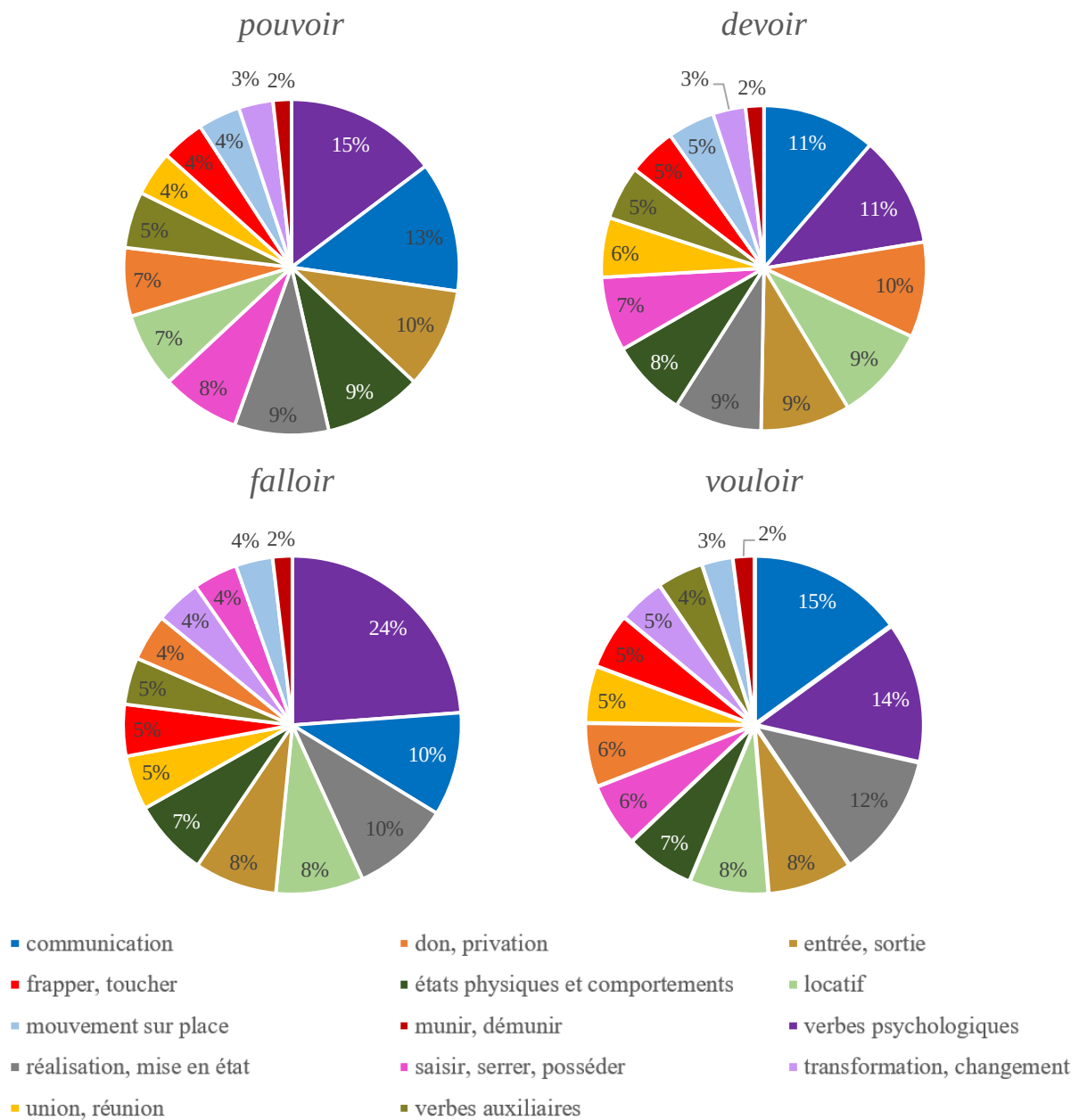


Figure 50 : Répartition des verbes à l'infinitif en position de complémentation verbale selon la classification de Dubois et Dubois-Charlier (1997) pour les formes lemmatisées des modaux dans le corpus LM.

Au-delà de ces deux premières catégories les plus représentées, les verbes divergent quelque peu dans la suite du classement : *pouvoir* attire ensuite des verbes de type *entrée/sortie* puis *d'états physiques/comportements*, *devoir* des verbes de type *don/privation* puis *locatifs*, *falloir* des verbes de *réalisation/mise en état* puis *locatifs* et enfin *vouloir*, des verbes de *réalisation/mise en état* puis de type *entrée/sortie*. A priori, il semble y avoir ainsi des ressemblances marquées entre les verbes mais également des divergences plus ou moins importantes au-delà des catégories les plus représentées.

Nous allons donc dans un premier temps nous pencher sur les verbes de communication qui représentent une catégorie importante pour ces quatre verbes en position de complémentation verbale. Nous consacrerons ensuite une section aux verbes psychologiques. Les autres associations seront étudiées pour *devoir* dans les sections suivantes car elles sont plus importantes pour les formes du présent de l'indicatif ainsi que pour celles du conditionnel présent.

7.2.3.2. Les complémentations verbales de type verbes de communication

Les verbes de communication représentent la catégorie la plus importante pour *devoir* et *vouloir* et la deuxième plus importante pour *pouvoir* et *falloir* en tant que lemmes dans le corpus LM. Parmi les 30 verbes ayant l'indice le plus haut dans cette catégorie en position de complémentation verbale, sept sont en commun pour les quatre modaux analysés ici :

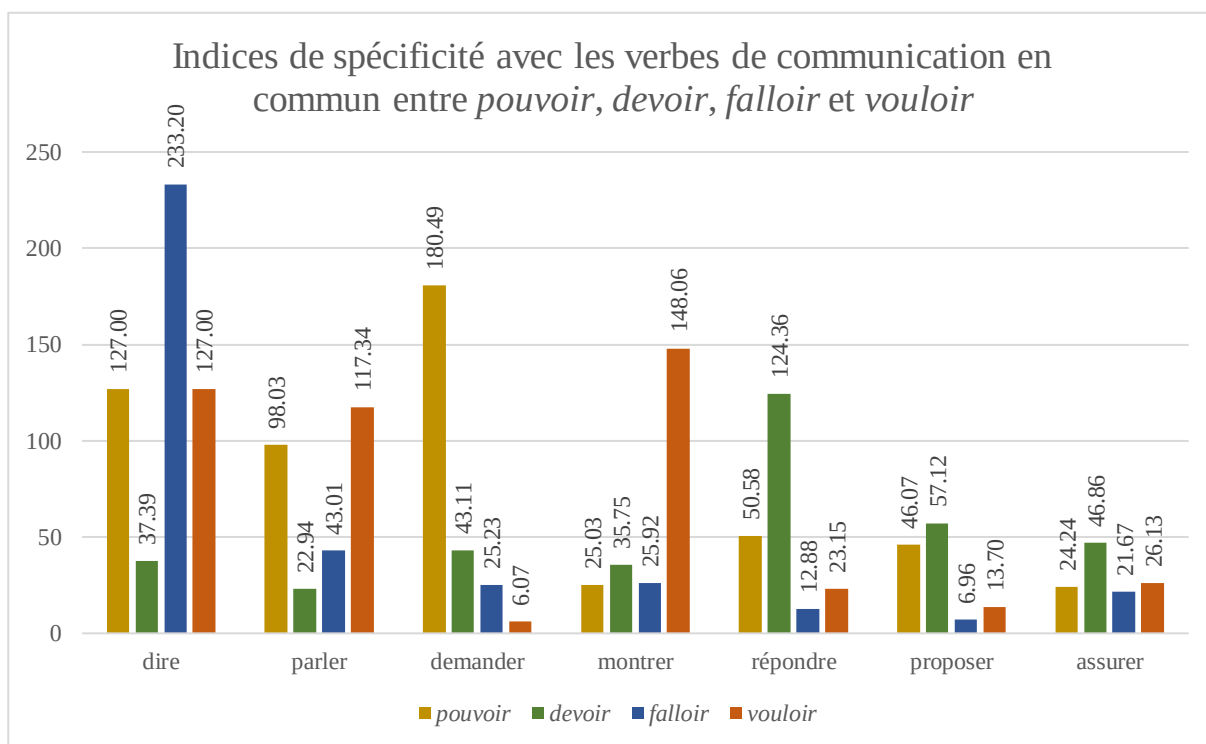


Figure 51 : Indice de spécificité pour les verbes de communication en commun pour les quatre modaux sous leur forme lemmatisée dans le corpus LM.

Le verbe *dire*, le plus prototypique des verbes de communication, est commun aux quatre modaux étudiés ici. L'indice le plus fort est celui avec *falloir* qui, nous le rappelons, se trouve grammaticalisé dans la locution *faut dire*. Cependant, le phénomène de grammaticalisation concerne également *vouloir* mais l'indice est ici moins important que pour *falloir*. *Pouvoir* a le même indice que *vouloir* alors qu'aucun phénomène de grammaticalisation n'a été attribué jusqu'à présent à ce verbe. *Devoir* présente ici l'indice le moins important avec *dire*.

Pour *falloir*, dans l'échantillon que nous avons observé, toutes les occurrences démontrent le phénomène de grammaticalisation :

189. C'est terrible, cette peur de perdre son boulot, chacun se barricade ! Il **faut dire** que ça valse, à France Télévisions : les responsables ne restent jamais suffisamment longtemps pour mener une politique cohérente, pour imprimer leur marque, leur vision (LM, 05.04.2010)
190. Les parlementaires et les policiers qui nous accompagnent ne sont pas très à l'aise. Il **faut dire** que la semaine dernière un de leurs collègues a reçu, ici même, un projectile sur la tête (dix points de suture). (LM, 12.04.2010)
191. Dans un acte que M. Cameron a qualifié de « désespéré », il a proposé à ses deux concurrents de cosigner une lettre adressée aux trois chaînes de télévision qui diffusent les débats entre les trois candidats. Alors que le troisième est prévu jeudi 29 avril, le premier ministre leur demandait de faire moins de publicité à l'événement, au motif qu'il occulterait des discussions plus fondamentales sur les programmes des candidats. Il **faut dire** que, jusqu'ici, l'exercice n'a pas réussi à cet homme peu télégénique. MM. Clegg et Cameron ont refusé. (LM, 29.04.2010)
192. Au 232, bordel de luxe de Melbourne, les passes avec les clients empruntent leur vocabulaire au cinéma. Il **faut dire** que, lors de ces « séances », au gré des fantasmes des clients, les six call-girls sortent le grand jeu en termes de scénario, d'interprétation et de costumes. (LM, 19.04.2010)

Falloir agit ainsi comme un marqueur évidentiel dans ce type de locution comme cela a été relevé par Pusch (2007) (cf. la section 4.4.5.3. de cette thèse), et donc comme un marqueur discursif selon les critères de Dostie (2004) puisqu'il a pour but de capter l'attention de l'interlocutrice sur l'énoncé (4.1.1.).

Concernant *vouloir*, les phénomènes de grammaticalisation sont évidemment nombreux avec cette association avec le verbe *dire*, et ce même avec une négation (195 et 196) ou avec un autre temps verbal que le présent (197) :

193. En Europe, on ignore ce que **veut dire** griot : pas seulement un conteur, mais tout à la fois le dépositaire de la mémoire de son peuple, mémoire uniquement orale, un maître de la parole (LM, 24.04.2010)
194. Ces rassemblements sont aussi une façon de trouver du plaisir en étant ensemble. Le problème, c'est qu'aujourd'hui, être ensemble, c'est picoler. Ça **veut dire** qu'on n'a pas réussi à mettre en place autre chose pour être ensemble sans être obligé de boire. (LM, 30.04.2010)

195. Ce projet connaît une déclinaison particulière à l'école Jean-Jaurès : les élèves y réalisent entièrement - ce qui ne **veut** pas **dire** seuls - le numéro de septembre 2010 du magazine de l'association, qui sera envoyé à ses 70 000 adhérents. (LM, 12.05.2010)
196. Nous continuons les discussions, il faut respecter la continuité de l'Etat. Poursuivre les négociations ne **veut** pas **dire** qu'elles vont aboutir. (LM, 15.05.2010)
197. Ces pages, poignantes d'authenticité, forment un témoignage sans équivalent sur ce que **voulait dire**, concrètement, aider les juifs sous l'Occupation. (LM, 09.04.2010)

Les occurrences de grammaticalisation pour *vouloir* + *dire* sont percevables lorsque nous pouvons considérer la locution comme un marqueur de reformulation (Franckel 2018). Au-delà de la reformulation, ces extraits de texte illustrent un effet de sens de *signification*. L'énoncé en 193, par exemple, présente une *signification* d'un terme sans que cela ne soit considéré comme une reformulation.

Dans certains exemples, *vouloir* est suivi du verbe *dire* mais n'est pas grammaticalisé :

198. « Quelles ont été vos indemnités ? - Nulles. - Ah bon, je croyais que c'était 350 000 francs ? - Euh, oui, c'est maladroite. Je **veux dire** que c'était juste le minimum légal. - Les pièces jaunes, en quelque sorte », dit le président. (LM, 24.04.2010)
199. Une porte-parole de la banque n'a pas **voulue dire** si ces anomalies étaient liées au client ou au conseiller financier. (LM, 05.04.2010)
200. A certains moments importants de l'histoire, l'opinion collective peut se tromper. La crise écologique me paraît relever de ces moments, tant l'ampleur de ses conséquences possibles serait grave. C'est tout ce que j'ai **voulue dire**, je ne voulais bien entendu pas stigmatiser tous ceux qui doutent. (LM, 05.04.2010)

Ici, on a bien une *intention/volonté* de dire quelque chose. Cette lecture est possible car les sujets sont ici des *animés*, ce qui confère à *vouloir* son effet de sens initial tiré directement de son NS. Rappelons que la locution *vouloir dire* se désémantise seulement si le sujet représente autre chose qu'un être *animé* (Franckel 2017).

Pour *pouvoir*, l'interprétation est, encore une fois, plus complexe à condenser étant donné que les effets de sens et les valeurs modales sont nombreuses pour ce verbe. Il semblerait que la *capacité* (201 et 202), la *permission* (203) ou encore la *possibilité* (204) soient toutes des interprétations possibles de cette association de *pouvoir* avec le verbe *dire* :

201. A Berlin, Xavier Villetard a rencontré des femmes qui témoignent. Le plus impressionnant est ce qu'elles ne disent pas, ne **peuvent** pas **dire**. « C'est enfermé dans mon cœur, voyez, là », dit une femme, geste à l'appui. (LM, 10.09.2010)
202. Nous en avons parlé. Je lui ai expliqué les choses, et il était tout à fait gentil. Voilà. Je ne **peux** pas vous **dire** mieux. Je me suis excusé auprès de lui en privé, et je le fais même en public à France Inter. (LM, 01.12.2010)
203. De son côté, sa femme voit cette thérapie comme « une pause où l'on **peut** se **dire** ce qui ne va pas, alors qu'on ne l'aurait jamais fait » (LM, 04.01.2010)
204. Mettra-t-il d'accord les deux tutelles ? Une réunion conjointe, prévue le 30 septembre, **pourrait** le **dire**. (LM, 21.09.2010)

En outre, les effets de sens que nous proposons ci-dessus sont contestables car le cotexte n'est pas assez explicite pour interpréter de manière tranchée le modal dans ces exemples.

Les interprétations sémantiques de *devoir* sont davantage aisées. Ainsi, de nombreuses fois c'est la valeur modale déontique qui ressort de ces exemples (205-208) et plus rarement la valeur modale épistémique (209) :

205. « Lundi dernier, j'avais deux places disponibles, elles ont été prises d'assaut. J'**ai dû dire** non 23 fois dans une soirée », comptabilise Joseph Jouffre. (LM, 30.11.2010)
206. Or GMC avait demandé que le texte soit signé par tous les délégués syndicaux. Cette condition n'étant pas remplie, le groupe **devra dire** dans les prochains jours s'il maintient son offre ou non. (LM, 26.07.2010)
207. Surtout, il a regretté d'avoir utilisé les termes « semblable à la Terre » (Earthlike) là où il **aurait dû dire** « de taille similaire à celle de la Terre » (Earth-sized). (LM, 04.12.2010)
208. Elle raconte la difficulté de de voir dire aux patients que leur chimiothérapie sera reportée de deux ou trois jours. (LM, 13.11.2010)
209. Toshiba lancerait des téléviseurs 3D visibles sans lunette ont créé un certain émoi chez ses concurrents qui ne proposent que des écrans 3D à regarder avec cet accessoire. Toshiba **devrait en dire** un peu plus au mois d'octobre. (LM, 07.09.2010)

La locution *je dois dire que* apparaît également dans ce corpus dans les extraits reprenant des dires oraux :

210. Mais, peu à peu, ma plume s'est installée dans cette langue mienne. Et je **dois dire** que, en fin de compte, c'est sans doute, de tous mes livres, celui que j'ai écrit avec le plus d'aisance. (LM, 16.04.2010)

211. Je **dois** d'abord **dire** que je suis un très gros consommateur de spectacles de théâtre et de danse. (LM, 29.03.2010)
212. « Satan, saloplaud, sais-tu au moins à quoi tient le succès de cette femme auprès des hommes ? » Là, je **dois dire** que j'étais bien embêté. Je ne pensais pas une seconde qu'il y eût besoin d'une raison. (LM, 10.02.2010)

Ici, la locution agit comme un « connecteur concessif » (Kronning 1988) comme nous l'avons relevé dans la partie 4.3.5.2. Il y a donc quelques occurrences de grammaticalisation pour *devoir* avec le verbe *dire* même si elles sont plus rares que pour les autres modaux que nous étudions ici, car ces exemples n'apparaissent que dans des transcriptions de et à la première personne à l'oral, occurrences habituellement peu fréquentes dans un corpus de presse.

En somme, les verbes modaux s'associent tous avec *dire* mais présentent également leurs particularités individuelles : *falloir* est toujours grammaticalisé avec ce verbe – du moins dans l'échantillon observé –, *vouloir* pas systématiquement – cela dépend du type de sujet grammatical qui le régit –, *pouvoir* présente des effets de sens très diversifiés et *devoir*, lorsqu'il n'est pas dans la locution *je dois dire que*, présente les valeurs modales épistémiques et déontiques classiques qu'on lui a toujours attribuées.

Pour avoir également un autre point de vue sur ces associations des modaux, nous allons également observer les occurrences possédant l'indice de spécificité les plus importants pour *pouvoir*, *devoir* et *vouloir* avec d'autres verbes de *communication* (pour *falloir* il s'agit de *dire* que nous venons d'étudier).

Pouvoir s'associe de manière importante avec le verbe *demande*r et celui-ci semble conférer au modal un effet de sens de *permission*, dans le sens d'*avoir le droit/la permission de demander* quelque chose :

213. Attention, passé 16 ans, les jeunes **peuvent** leur **demande**r des comptes et les parents n'ont plus le droit de piocher dans leur épargne. (LM, 03.04.2010)
214. Après le décès d'un salarié, le conjoint survivant **peut demande**r une pension de réversion, même s'il s'agit d'un ex-conjoint qui s'est remarié. (LM, 08.05.2010)

Ici, c'est donc bien le poids sémantique de *demande*r qui teinte l'interprétation de *pouvoir* de manière unanime dans l'échantillon que nous avons analysé. Nous pouvons ici à nouveau comprendre cette valeur modale déontique par inférence.

Devoir, quant à lui, possède l'indice le plus haut avec le verbe *répondre* qui lui confère la valeur modale déontique même lorsque le sujet n'est pas *animé* comme en 217 :

215. Les questions auxquelles **devront répondre** les experts sont précisées dans l'ordonnance de la juge. (LM, 12.01.2010)
216. Après chaque saynète, « tu » **dois** donc **répondre** à des QCM, genre phrases à trous qu'il faut remplir dans les cahiers de vacances. (LM, 09.07.2010)
217. Pour Patrick Garcia, « l'apprentissage de l'histoire, depuis la fin du XIXe siècle, **doit répondre** à une triple finalité : morale, intégratrice et civique. (LM, 13.09.2010)

Dans 215, la valeur modale déontique peut être confirmée, non seulement par le temps verbal du futur – qui est très souvent un indicateur de ce type de modalité avec *devoir* comme le signalait Huot (1974) (4.3.8.2.) – mais aussi par inférence car nous avons bel et bien une figure d'autorité (la *juge*) qui ordonne ici. Dans 216, l'effet de sens d'*obligation* peut également être inféré voire renforcé grâce à l'autre modal *falloir* qui figure dans la suite de l'énoncé.

Vouloir s'associe de manière la plus significative avec le verbe *montrer*. Dans les occurrences que nous avons observées, nous avons pu noter qu'il s'agit d'une expression de son NS d'*intention/volonté* :

218. Le choix des images était évidemment un choix politique, un choix idéologique. Il n'y avait là aucune objectivité. La télévision montrait ce qu'elle **voulait montrer**. (LM, 03.09.2010)
219. Nous **avons voulu montrer** que les effets de l'éducation sur les sociétés s'expliquaient, pour une part, par les fonctions attribuées à l'école. (LM, 26.08.2010)
220. « Il est brillant et n'a de cesse de **vouloir le montrer**. Il est séduisant et veut constamment user de son charme » (LM, 21.04.2010)
221. C'est l'entente cordiale, où chacun **veut montrer** qu'il est le plus fort, pas non plus l'amour fou. (LM, 18.11.2010)

Étant donné les associations importantes et récurrentes entre les verbes modaux et les verbes de *communication*, nous les comparons dans une AFC. Cette approche permet de voir les rapprochements et éloignements entre les modaux par rapport à ce groupe en particulier :

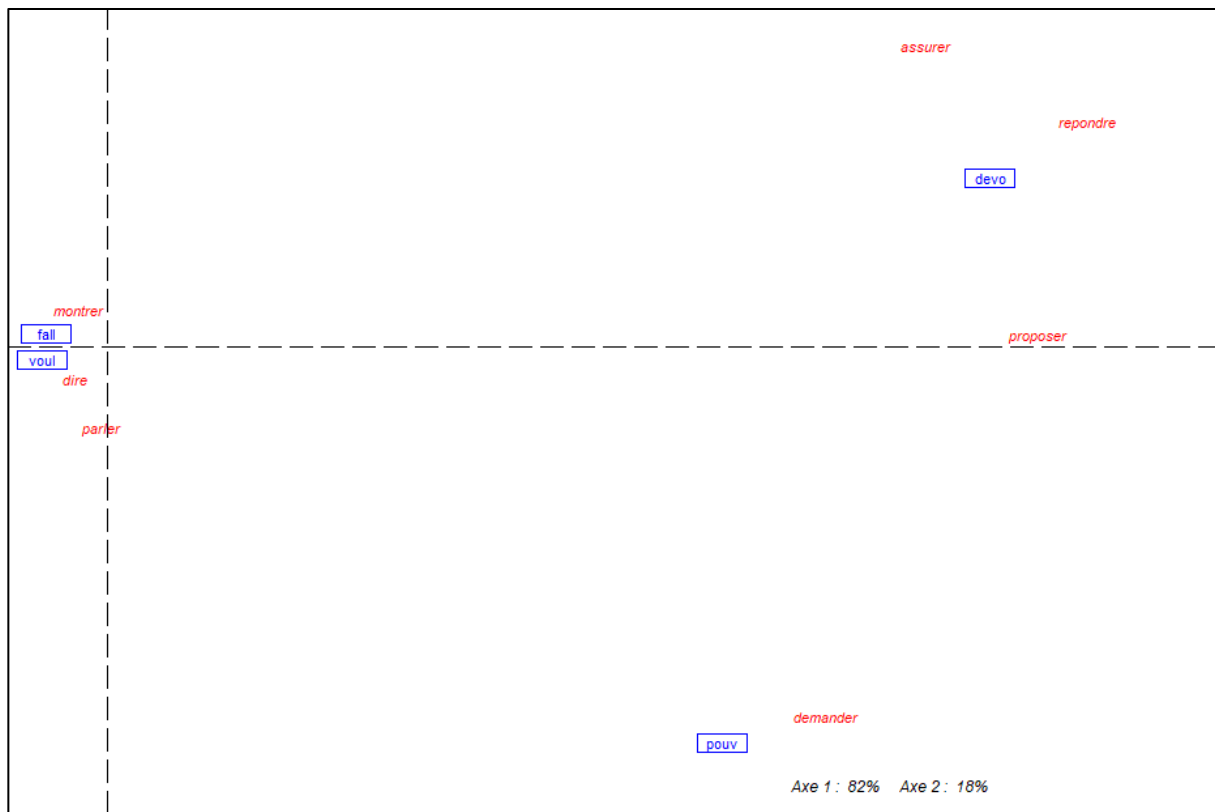


Figure 52 : AFC des modaux sous leur forme lemmatisée avec les verbes de communication en commun en position de complémentation verbale dans les 30 indices les plus élevés dans LM.

Nous retrouvons ici une opposition importante sur l'axe 1, entre *devoir* et *pouvoir* à droite de l'axe et *falloir* et *vouloir* à gauche, et proches du centre. Ce sont les deux verbes les plus grammaticalisés de notre paradigme. Ils sont d'ailleurs très proches de *dire* dans cette figure 54. C'est donc en grande partie les associations désémantisées avec ce verbe qui les placent sur ce point-là du graphique. Ils sont également proches de *montrer* (qui avait l'indice de spécificité le plus important avec *vouloir*) et de *parler* (qui comme *dire*, semble être un verbe de communication générique).

Cependant, *pouvoir* et *devoir* s'opposent sur l'axe 2. *Pouvoir* est isolé en bas à droite du graphique avec *demande* (dont l'indice de spécificité était le plus important). Ceci confirme ainsi que c'est par cette association qu'il se démarque des autres verbes. *Devoir*, dans le cadran en haut à droite, est entouré des verbes *assurer*, *répondre* et *proposer*. Ces verbes sont ici plus chargés sémantiquement que ceux proches du centre, ils impliquent « un changement d'attitude mentale, le destinataire est aussi expérimenté » (Abeillé & Koenig 2021b : 199). *Devoir* semble ainsi modaliser des verbes de communication « forts », alors que *pouvoir* adoucit une action de *demande* et *falloir* et *vouloir* se grammaticalisent de manière importante avec le verbe prototypique de cette catégorie et sémantiquement neutre, *dire*.

Ainsi, même si tous ces verbes modaux présentent une attraction particulière avec les verbes de *communication* et que certains de ces verbes sont des cooccurrents communs des modaux, des singularités peuvent être soulevées pour chacun de ces modaux lorsque nous étudions de plus près les indices de spécificité et leurs positions sur l'AFC.

7.2.3.3. Les complémentations verbales de type verbe psychologique

Concernant les verbes classés comme *psychologiques*, seulement quatre sont en commun pour les quatre modaux étudiés, parmi les 30 associations avec le plus grand indice de spécificité de cette catégorie :

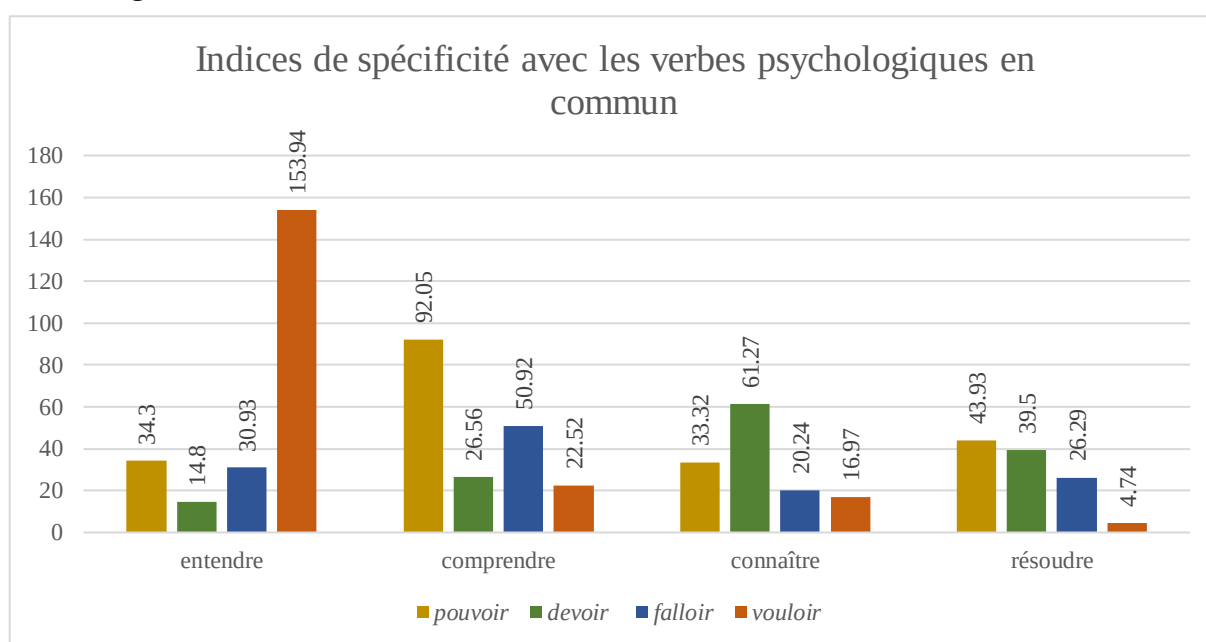


Figure 53 : Indice de spécificité pour les verbes psychologiques en commun en position de complémentation verbale pour les quatre modaux étudiés sous leur forme lemmatisée dans le corpus LM.

Alors même que tous les modaux sont attirés de manière significative avec ces quatre verbes, nous remarquons que les indices sont relativement différents selon les associations. *Pouvoir* et *falloir* sont majoritairement attirés par *comprendre*, *devoir* par *connaître* et *vouloir* par *entendre*. Ce sont donc ces associations que nous allons observer en premier.

Concernant les associations de *pouvoir* et *falloir* avec *comprendre*, nous pouvons relever que la plupart permet de mettre en avant un « fait général à comprendre ». Pour *pouvoir*, l'effet de sens oscille à nouveau entre la *possibilité* et la *capacité*, mais quelle que soit l'interprétation adoptée, c'est une *généralité* qui est exprimée à l'aide du pronom *on* qui précède cette association :

222. On discutait avec les Américains depuis huit ans et on négociait le traité depuis deux ans. On **pouvait comprendre** un changement de paramètres, mais pas la méthode employée. (LM, 09.04.2010)
223. Ce qui définit le philosophe, c'est toujours l'idée que l'on **peut comprendre** l'autre, que l'on peut comprendre l'adversaire. (LM, 01.10.2010)
224. La majorité évite soigneusement le sujet. On **peut comprendre** sa gêne quand on sait le coût, 3 milliards d'euros au total en 2010, des exonérations supplémentaires des droits de succession votées notamment dans le paquet fiscal, en 2007. (LM, 10.12.2010)

Cela rappelle les remarques de Sitri (2013) reprises par Née et al. (2016) :

Sitri montre que, dans le corpus, un emploi combinant la capacité et la sporadicité est caractéristique. Cet emploi prend un sens spécifique que l'on peut paraphraser par « le procès a eu lieu et c'est notable ». Cette valeur de sporadicité-notabilité, note l'auteur, peut venir illustrer un énoncé généralisant, enchainement susceptible de se figer dans la routine « et peut + INF ». (Née et al. 2016 : 83)

Dans notre corpus, il ne s'agit pas de cette routine en particulier, mais la construction *on + peut + verbe communication* semble également montrer la notabilité d'un énoncé généralisant.

La généralité se retrouve aussi avec les formes impersonnelles de *falloir*, mais dans les exemples suivants, il semble *nécessaire* de *comprendre*. L'injonction n'est ici plus suggérée mais bien exprimée à travers la valeur modale déontique :

225. Que cette planète intrigue, fascine ou scandalise, il **faut** d'abord **comprendre** comment elle tourne. (LM, 13.03.2010)
226. Lorsqu'ils écrivent qu' « on ne peut pas tout décider depuis les bureaux à Paris », il **faut comprendre** depuis l'Elysée. (LM, 06.09.2010)
227. Si votre interlocuteur vous lance un « Je serai à la datcha ce dimanche ! », il **faut comprendre** : « Ne cherchez surtout pas à me joindre ! » (LM, 14.05.2010)

Pouvoir et *falloir* se ressemblent donc sur ce point avec leur association importante avec le verbe *comprendre* : c'est une injonction plus ou moins forte adressée à la lectrice du journal qui *devrait* ou *doit* comprendre ce qu'il en est, ou, du moins, remarquer la notabilité d'un fait généralisant.

Devoir s'associe de manière importante avec le verbe *connaître*. Cependant, ces associations montrent un usage de *connaître* en tant que verbe d'*expérience* plutôt que de *pensée* :

228. Selon la société d'études Eurostaf, le marché mondial du luxe **devrait connaître** en 2010 une croissance comprise entre 2 % et 4 % hors effet de change. (LM, 06.02.2010)
229. D'autres dossiers aussi techniques **devraient connaître** des avancées plus aisées, comme la définition de cadres institutionnels facilitant l'élaboration de politiques d'adaptation au changement climatique ou le transfert de technologies entre pays riches et pays en développement, même si la question des droits de propriété intellectuelle est loin d'être tranchée. (LM, 29.11.2010)
230. Amir Kasab, qui encourt la peine de mort, **devait connaître** sa condamnation mardi. (LM, 05.05.2010)

Connaître est toutefois toujours considéré comme un verbe psychologique chez Dubois et Dubois-Charlier (1997). Cette catégorie comprend, en effet, des verbes d'*activité mentale* dans le sens large du terme, donc non seulement relatif aux *croyances* et *pensées* mais aussi aux *expériences sensorielles* et aux *connaissances*. Toutefois, ces exemples montrent à nouveau la tendance de *devoir* à s'associer de manière plus significative avec des verbes dénotant des actions plus concrètes¹⁸³ comme nous l'avons relevé avec l'AFC des verbes de *communication* (figure 53).

Le verbe *entendre* est donc également une expérience sensorielle, c'est pourquoi il apparaît parmi les verbes *psychologiques*. Il présente avec *vouloir* un indice de spécificité important. Ces cooccurrences montrent des usages dans lesquels *vouloir* se grammaticalise quelque peu car il est suivi par des locutions figées comme *en entendre parler* (231 et 232) et *à qui veut l'entendre* (233 et 234) :

231. La semaine dernière, le ministre de la culture, Frédéric Mitterrand, avait évoqué l'éventualité d'une hausse de la redevance, une solution dont M. Copé ne **veut** pas **entendre** parler. (LM, 08.04.2010)
232. « Ce serait plus simple, plus efficace et plus équitable. » Mais l'Elysée ne **veut** pas **entendre** parler, évoquant des « mesures aveugles » et « pas très courageuses ». (LM, 28.08.2010)

¹⁸³ Nous comprenons cette différenciation dans le sens de Sitri (2013) – qui nous rappelle Vendler (1957) – lorsqu'elle écrit : « les prédicats ainsi modalisés par *pouvoir* décrivent en grande majorité des activités (*exprimer ses désirs*) et non des accomplissements (*élaborer* par exemple) et que l'événement en question opère une transformation du sujet et non de l'objet » (Sitri 2013 : 83). Il semblerait que ce soit également le cas ici lorsqu'on différencie les verbes plus attirés par *pouvoir* que par *devoir*. Nous considérons ainsi les accomplissements comme pouvant être représentés par des verbes dénotant des actions plus « actives » et « concrètes ».

233. Eclaboussé par l'affaire, Eugène Caselli assure désormais à qui **veut** l'**entendre** qu'il répondra à toutes les questions que lui posera la justice. (LM, 17.12.2010)
234. Les Russes disent à qui **veut** l'**entendre** que si nous allons jusqu'à la 4e phase du déploiement du système antimissiles, ils se retireront du traité. (LM, 23.12.2010)

C'est ce type de signification qui se lit le plus fréquemment pour cette association même si le sens « plein » de *vouloir* apparaît également avec *entendre* :

235. Dominique de Villepin **veut entendre** ces Français qui souffrent. (LM, 21.01.2010)
236. Nicolas Sarkozy et François Fillon ne **veulent pas entendre** ce que disent les Français. (LM, 20.03.2010)

En somme, *vouloir* présente à nouveau de nombreux usages où il se désémantise dans les cooccurrences avec un indice de spécificité important.

7.2.3.4. Les répartitions des verbes en cooccurrences significatives pour les formes au présent de l'indicatif

En observant les répartitions des catégories sémantiques des complémentations verbales qui apparaissent de manière significative pour les modaux lorsqu'ils sont conjugués au présent de l'indicatif dans le corpus LM, nous remarquons qu'il y a de légères différences avec ce que nous avons observé pour les formes lemmatisées :

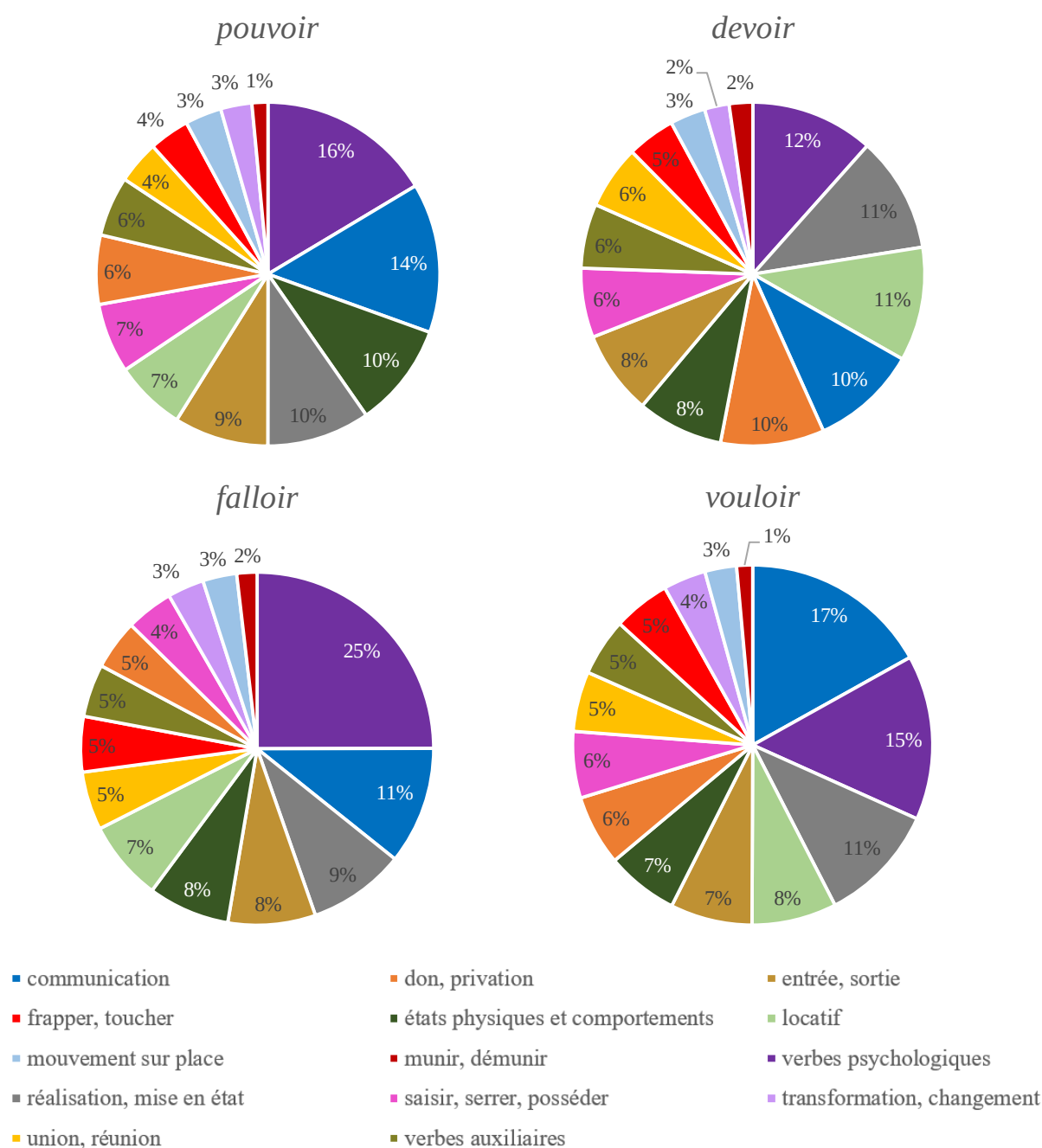


Figure 54 : Répartition des verbes à l'infinitif en position de complémentation verbale selon la classification de Dubois et Dubois-Charlier (1997) pour les formes du présent de l'indicatif des modaux dans le corpus LM.

Pouvoir, *devoir* et *falloir* attirent majoritairement des verbes *psychologiques*. *Vouloir* préfère de peu les verbes de communication mais les verbes *psychologiques* arrivent en deuxième position. *Pouvoir* et *falloir* ont les verbes de *communication* en deuxième position également. Sur ce point, *devoir* diverge fortement lorsqu'il est employé au présent de l'indicatif : avant les verbes de *communication*, ce sont les verbes de *réalisation/mise en état* ou encore les verbes *locatifs* qui occupent une part plus importante parmi les verbes en position de complémentation verbal en cooccurrence significative.

Nous allons donc voir de plus près ces ressemblances et ces divergences en contexte. Pour *pouvoir* et *falloir*, nous observerons leurs cooccurrences avec les verbes *psychologiques*, pour *devoir*, ses usages avec les verbes de *réalisation/mise en état* et les *locatifs* et pour *vouloir*, nous jetterons un coup d’œil à ses associations avec les verbes de *communication*.

7.2.3.5. Les complémentations verbales de type verbes psychologiques avec *pouvoir* et *falloir* au présent de l’indicatif

Pouvoir et *falloir* partagent énormément de verbes en commun lorsqu’ils sont au présent de l’indicatif avec la catégorie des verbes *psychologiques* puisque dans les 30 verbes ayant l’indice de spécificité le plus haut, 12 se trouvent être les mêmes :

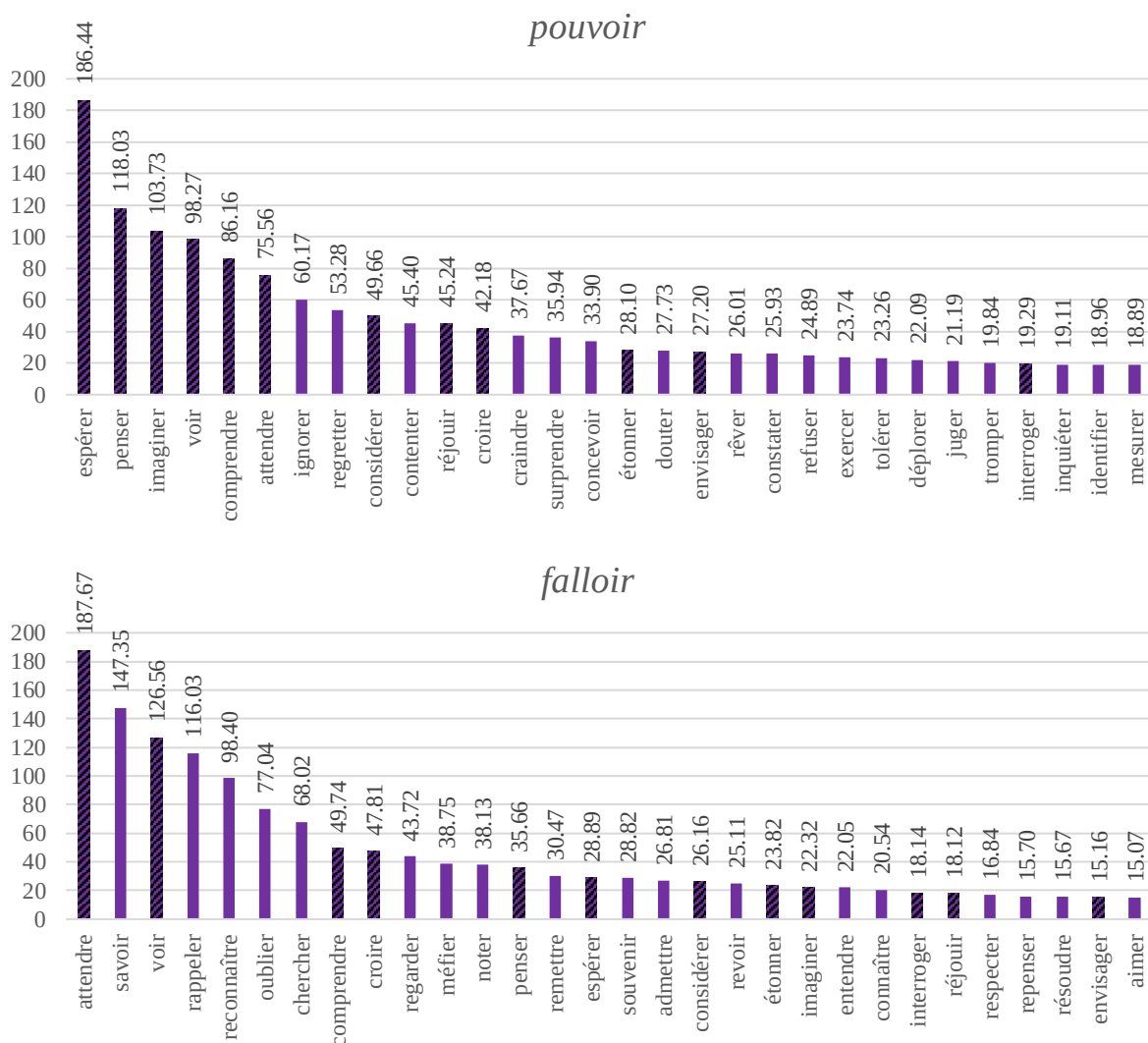


Figure 55 : Les 30 verbes de type psychologiques en position de complémentation verbale avec le plus grand indice de spécificité pour *pouvoir* et *falloir* au présent de l’indicatif dans LM

Pour *falloir*, nous retrouvons indéniablement la locution *faut croire* qui a été qualifié de marqueur évidentiel (Rossari 2012) :

237. J'avais une chance sur deux de naître du bon côté : **faut croire** que je suis née sous une bonne étoile. (LM, 08.03.2010)

L'omission du pronom *il* montre ici bien la grammaticalisation de cette expression comme l'avaient soulevé Pusch (2007) ou encore Valiukienė (2022) (4.4.5.).

Falloir semble également se grammaticaliser avec d'autres verbes *psychologiques*, notamment *voir* ou encore *imaginer* comme le relevait Valiukienė (2022) (4.4.5.1.) car le NS de *nécessité* ou encore de *contrainte* est plus difficile à récupérer dans ce type d'usage :

238. Pour le moment, la séquence Aubry a été plutôt positive, mais il **faut voir** si cet avantage se maintient dans la durée. (LM, 27.05.2010)
239. La future présidente aurait pu lui proposer un secrétariat d'Etat au tourisme ou à la vie associative. Mais non : premier ministre d'un coup ! Il **faut les imaginer** sur un tandem : elle, à l'avant, tenant le guidon, et lui, à l'arrière, suant et soufflant, en train de pédaler de toutes ses forces. (LM, 02.12.2010)

Ces locutions agissent comme des marqueurs de focalisation comme les avait dénommés Valiukienė (2022). Elles ont pour but d'attirer l'attention des lectrices sur un phénomène en particulier. Au vu de cette fonction, nous pouvons donc aussi le rapprocher de la notion de marqueur discursif comme définie par Dostie (2004) (*cf.* 4.1.1.).

Cependant, la grammaticalisation n'est pas automatique ou pas encore au dernier stade avec les verbes psychologiques, car *falloir* peut également arborer un effet de sens ou une valeur modale plus proche de son NS. Dans les exemples suivants, il y a bien une *contrainte* ou une *nécessité* générale qui est percevable :

240. On va le faire mais quand, ça, je ne sais pas, il **faut attendre** que l'eau redescende, les autorités disent dans un mois peut-être. (LM, 27.08.2010)
241. Parce que, pour juger, il **faut comprendre**. (LM, 09.08.2010)
242. Et on insiste sur le fait qu'il **faut considérer** l'ensemble du cycle avant de se prononcer, et non isoler telle ou telle année. (LM, 29.05.2010)

Avec *savoir*, qui est le verbe avec le plus gros indice de spécificité pour *falloir*, il nous semble que l'effet de sens d'*obligation* du modal est toujours perceptible :

243. François Bayle a conservé le souvenir des bruits de la nature, des animaux qu'on entend avant de les voir, bruits inquiétants ou rassurants qu'il **faut savoir** déchiffrer pour se préserver ou qu'on peut écouter comme une musique. (LM, 05.04.2010)
244. La laïcité ne doit pas rester une idéologie. Il **faut savoir** être pragmatique pour permettre des accommodements, sinon on se tire une balle dans le pied face à cette jeunesse musulmane, qui se sent souvent stigmatisée. (LM, 05.04.2010)
245. Le milieu de l'art exige de ses jeunes recrues toujours plus de dossiers, qu'il **faut savoir** mettre en forme. (LM, 10.04.2010)
246. Celui qui avait théorisé, avant son élection de 2007, que pour réformer il **faut savoir** mobiliser son camp (la droite), voudrait maintenant réformer en rassurant d'abord ses opposants (la gauche). (LM, 22.05.2010)
247. Il **faut savoir** dire non quand les droits humains sont bafoués. (LM, 11.06.2010)

Pour les exemples 243 et 244, nous pouvons remarquer ici qu'il s'agit d'une *contrainte tournée vers un objectif ou un but* si on reprend la terminologie de Caron et Caron-Pargue (2003) (4.4.3.). Cette interprétation est surtout possible grâce à la préposition *pour* qui suit la construction. Pour 245, 246 et 247, nous avons davantage à faire à des exemples dans lesquels *savoir* est un verbe modal dénotant une *capacité* plutôt qu'une action *psychologique*. Cela ne change toutefois pas l'interprétation de *falloir* qui semble ici dénoter une *obligation générale*. C'est sans doute dans ces usages de *contrainte* ou *nécessité générale* que *falloir* se rapproche de *pouvoir*. En effet, *pouvoir* semble être utilisé avec les verbes *psychologiques* pour exposer des faits généraux, comme nous le susmentionnions. Cette généralité ressort en particulier avec le pronom *on* dans les exemples que nous avons pu relever :

248. Voilà le genre de livre dont on **peut s'étonner** qu'il ait été publié. (...) L'écriture est si touffue, si prétentieuse et si truffée d'allusions littéraires que la lecture en est difficile. (LM, 29.04.2010)
249. Tout en rappelant, dans sa réponse à la question sur le préservatif, la théorie ABC- Abstinence, Be faithful (fidélité), Condom (préservatif)-, Benoît XVI est de fait le premier pape à reprendre à son compte le principe « du moindre mal » sur ce sujet. On **peut y voir** une forme de pragmatisme : cette approche est mise en avant depuis des années par des responsables religieux confrontés à des malades du sida, qui défendent l'idée que transmettre la mort est pire que transgresser un interdit. (LM, 23.11.2010)

L'association *on + peut + verbe psychologique* semble agir comme *on + peut + verbe communication*, et attire ainsi l'attention de la lectrice (ou de l'interlocutrice),

sur un fait notable général. *Pouvoir* le fait cependant de manière plus « douce » que *falloir* au vu de leurs NS distincts. Le fait semble encore plus « notable » lorsqu'il est mis en avant avec *falloir*. Toutefois, étant donné que *pouvoir* est précédé par *on* dans ces constructions et que *falloir* est par essence un verbe impersonnel, ils permettent tous deux de mettre en avant l'importance d'un fait pour la lectrice ou l'interlocutrice tout en maintenant la face de celle-ci (Brown & Levinson 1987 : 70). En effet, il serait plus néfaste pour la face d'écrire *vous devez y voir une forme de pragmatisme*, par exemple.

Cependant, lorsque *pouvoir* est précédé par un autre sujet que le pronom *on*, nous retrouvons facilement un des effets de sens déjà attesté, comme la *capacité* dans l'exemple suivant :

250. Cette petite révolution est le fruit des progrès de l'espérance de vie, qui augmente désormais de trois mois par an : un enfant né en 2010 **peut espérer** vivre jusqu'à plus de 80 ans, contre seulement 66 ans en 1950. (LM, 16.02.2010)

En somme, dans certains cas, le NS tend à disparaître lorsque *pouvoir* et *falloir* sont associés à des verbes *psychologiques* en position de complémentation verbale. Nous pourrions ainsi admettre qu'il s'agit ici de phases précoces de grammaticalisation.

7.2.3.6. Les complémentations verbales de type verbes réalisation/mise en état et locatifs avec devoir au présent de l'indicatif

Devoir s'associe de manière importante avec les verbes *réalisation/mise en état* ainsi que les verbes de type *locatif* dans le corpus LM lorsqu'il est conjugué au présent de l'indicatif. Ces verbes sont en somme assez similaires puisqu'ils décrivent une action concrète :

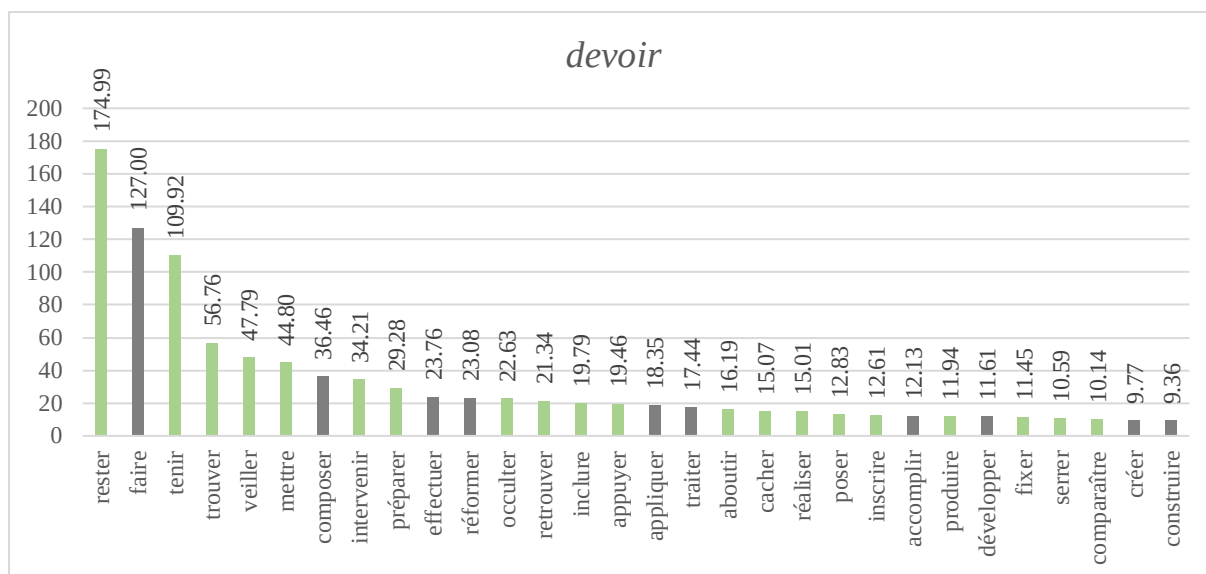


Figure 56 : Les 30 verbes de type locatif ou de réalisation/mise en état en position de complémentation verbale avec le plus grand indice de spécificité pour *devoir* au présent de l'indicatif dans LM.

En observant un échantillon de ces associations, *devoir* dénote dans la grande majorité des cas une modalité déontique. Ces exemples montrent qu'il y a une nécessité que quelque chose doit être réalisée dans un certain but :

251. Si l'on ne veut pas que les dettes souveraines des pays riches deviennent des chiffons de papier, ceux-ci **doivent mettre** en place des projets crédibles de réduction des dépenses. (LM, 23.02.2010)
252. Pour lancer la TMP, les seize chaînes autorisées par le Conseil supérieur de l'audiovisuel (CSA) **doivent créer** une société commune avant juin pour diffuser les programmes. (LM, 23.04.2010)
253. En 1962, le Vatican a consacré une longue directive au traitement de ce crime-là. Ce qui frappe, c'est qu'il **doit rester** secret. (LM, 26.03.2010)
254. A ses côtés, Bobby reprend goût à la vie mais, très vite, **doit se préparer** au pire... (LM, 15.03.2010)

Toutefois, cette valeur modale déontique ne résulte pas d'une analyse automatique de cette association. En effet, lorsqu'une date est mentionnée dans l'énoncé, le verbe *devoir* acquiert

souvent une nuance de sens prédictif par inférence, ce qui se rapproche davantage d'une valeur modale épistémique, avec degré de certitude élevé :

255. Elle a affirmé se préparer à des « fraudes massives » lors de ce scrutin, dont le premier tour **doit se tenir** dimanche 28 novembre. (LM, 27.11.2010)
256. Le contentieux franco-nigérien s'était finalement réglé entre le président français, Nicolas Sarkozy, et Mamadou Tandja, ouvrant la voie à la signature, en janvier 2009, d'une convention minière pour l'exploitation de la mine d'Imouraren, qui **doit produire** sa première tonne d'uranium en 2013. (LM, 18.09.2010)

En somme, il semble qu'avec des verbes dénotant une action plus tangible que des verbes de *communication*, *devoir* au présent permet d'exprimer une *nécessité* que cette action soit réalisée. Toutefois, celle-ci se trouve nuancée avec le co(n)texte général de l'énoncé : lorsqu'il comprend la mention d'une date, nous avons plutôt à faire à une *prédiction* qui, par nature, est moins certaine qu'une *nécessité* et se rattache ainsi davantage à la catégorie des valeurs modales épistémiques.

7.2.3.7. Les complémentations verbales de type verbes de communication avec vouloir au présent de l'indicatif

Vouloir s'associe de manière importante aux verbes de *communication*, ce qui, sans surprise, est en partie dû à la locution *vouloir dire* qui semble être très importante au présent :

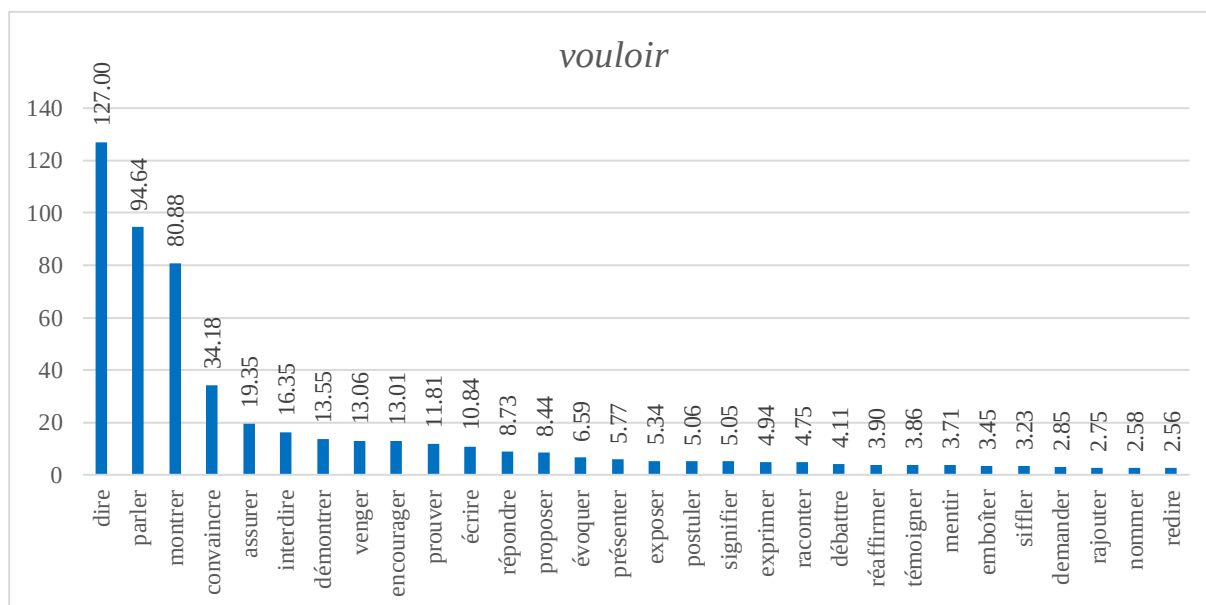


Figure 57 : Les 30 verbes de type communication en position de complémentation verbale avec le plus grand indice de spécificité pour vouloir au présent de l'indicatif dans LM.

Vouloir acquiert la fonction de *reformulation* – paraphrasable par le verbe *signifier* – lorsqu’il est associé à *dire* :

- 257. Car un parti, cela **veut dire** des moyens, une structure, un soutien. (LM, 11.02.2010)
- 258. Surtout, selon l’expression andalouse, Ramon Lopez est « una buena persona », ce qui ne **veut dire** ni un brave type ni une bonne personne, mais un être attentif, généreux. (LM, 30.08.2010)
- 259. Coupe du monde en Afrique ne **veut** pas **dire** Coupe du monde de l’Afrique... (LM, 15.06.2010)
- 260. « Nous ne devrions pas effrayer les chevaux, ce qui **veut dire** ne pas effrayer les Français - et garder les Russes et les Chinois à bord. » (LM, 23.12.2010)

Il existe des cas où la locution *veut dire* n’est pas figée comme nous pouvons le montrer dans les exemples suivants :

- 261. Ce n’est pas le mot qui compte, c’est ce que la personne **veut dire**. (LM, 28.07.2010)
- 262. Aujourd’hui, elle **veut dire** aux couples que ce n’est pas une nécessité d’avoir des enfants pour être heureux, que ça n’est pas non plus une obligation. (LM, 13.12.2010)
- 263. Si l’on **veut dire** trois mots de la naïveté d’Alexandre Romanès, on jouera aux devinettes. (LM, 04.06.2010)
- 264. En faisant respecter à tout prix « son » calendrier à l’Assemblée, le gouvernement **veut dire** aux salariés « la réforme est passée ! » (LM, 21.09.2010)
- 265. Maguy Marin sait ce qu’elle **veut dire** et trouve les moyens pour faire passer le message sans virer simpliste ni démagog. (LM, 16.09.2010)

Ces cas ne concernent cependant que 2,25% des cooccurrences de *vouloir* au présent de l’indicatif avec *dire*. Et, comme nous l’avons relevé, cette interprétation est possible seulement lorsque le sujet grammatical est représenté par un *animé*.

Lorsque *vouloir* est conjugué au pluriel et suivi par *dire*, il peut également se grammaticaliser (266 et 267). Nous ne pouvons, cependant, ici pas tirer de conclusion sur la fréquence de cette grammaticalisation lorsque le verbe est au pluriel car nous n’avons que trois occurrences de *veulent dire* dans LM dont une seulement n’est pas figée (268) :

- 266. Ceux qui voyaient dans la chute du mur de Berlin l’aurore d’une démocratisation heureuse et sans heurts ont tout simplement cessé de réfléchir à ce que les droits de l’homme **veulent dire**. (LM, 09.10.2010)

267. Que veut dire la notion de liberté, où sont ses limites et celles de la presse ? Que **veulent dire** Parlement et élections ? (LM, 02.03.2010)
268. Dans la salle des témoins, ils sont nombreux à attendre, membres du Rotary, relations professionnelles, compagnons de golf, tous membres du comité de soutien qui s'est créé autour de Jean-Michel Bissonnet et qui **veulent dire**, à leur tour, combien la justice se fourvoie en accusant leur ami. (LM, 04.10.2010)

A nouveau, si le sujet est un *animé*, nous pouvons aisément retrouver le NS d'*intention/volonté* pour *vouloir* qui ici détient un plus grand poids sur son interprétation sémantique que le verbe qui suit. C'est d'ailleurs les sujets *animés* que l'on retrouve avec les autres verbes de communication auxquels *vouloir* s'associe fréquemment :

269. Je **veux** lui **montrer** que je suis toujours en vie. (LM, 20.07.2010)
270. Si nous **voulons exprimer** la diversité de notre électorat, il faut que ces courants puissent s'exprimer. (LM, 01.09.2010)
271. En matière de réforme des retraites, rien n'est dit sur le niveau de pension, le niveau de vie que l'on **veut assurer** aux retraités. (LM, 10.05.2010)
272. Mais au « cérémonial de la violence », elle **veut répondre** par l'espoir, et la « fraternité de la parole ». (LM, 05.03.2010)

Il est ainsi difficile de dissocier ici les analyses de la complémentation verbale sous forme de verbe à l'infinitif de celles que nous avons réalisées avec les sujets. Ce sont bien les combinaisons de ces deux facteurs qui permettent une grammaticalisation de *vouloir* ou non.

7.2.3.8. Les répartitions des verbes en cooccurrences significatives pour les formes au conditionnel présent

Au conditionnel présent, nous avons, une nouvelle fois, d'autres profils combinatoires qui se dessinent pour les verbes :

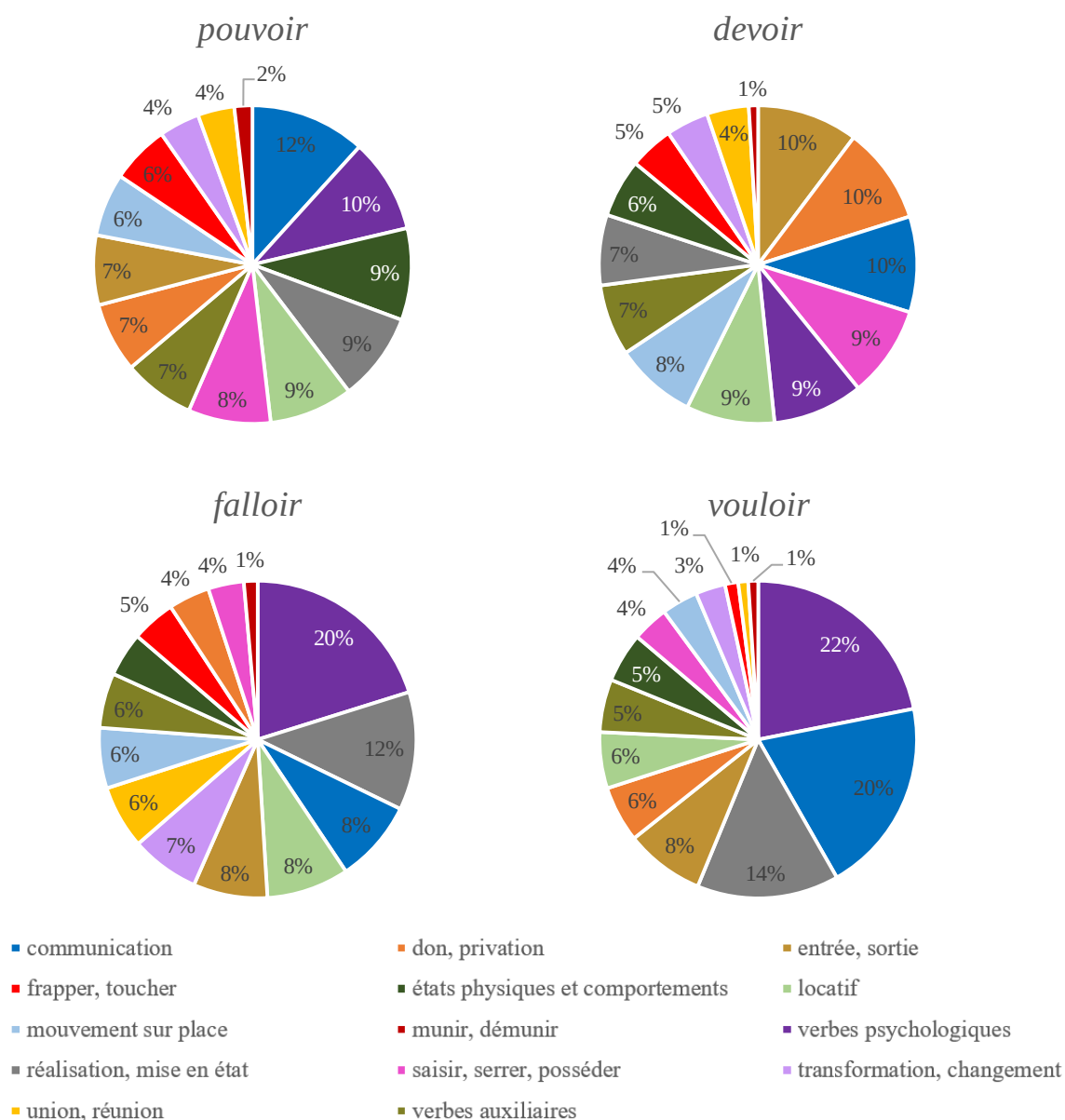


Figure 58 : Répartition des verbes à l'infinitif en position de complément verbale selon la classification de Dubois et Dubois-Charlier (1997) pour les formes du conditionnel présent des modaux dans le corpus LM.

Devoir est majoritairement attiré par des verbes de type *entrée/sortie*, *don/privation* puis seulement après par les verbes de *communication*. En somme, la répartition pour le conditionnel ne fait pas ressortir une catégorie de manière prédominante par rapport à une autre. C'est également le cas pour *pouvoir* dont les proportions des catégories sont tous très proches. Toutefois, ce sont à nouveau les verbes de *communication* et les verbes *psychologiques* qui arrivent en tête de liste, puis les verbes d'*états physiques/comportements*. Enfin, *falloir* et *vouloir* montrent une prédominance pour les verbes *psychologiques* et les verbes de *communication*.

Au vu des disparités que les verbes présentent ici dans leurs profils combinatoires lorsqu'ils sont au conditionnel présent, nous allons nous pencher des catégories quelques peu différentes pour chacun de ces modaux. Pour *pouvoir* et *vouloir*, nous observerons les cooccurrences avec les verbes de *communication* et verbes *psychologiques*, pour *devoir*, les trois catégories citées qui arrivent en tête du graphique et pour *falloir* les verbes *psychologiques*.

7.2.3.9. Les complémentations verbales de type verbe de communication et psychologiques avec pouvoir et vouloir au conditionnel présent

Pouvoir et *vouloir* se combinent tous deux de manière privilégiée avec les verbes de *communication* et ceux classés comme *psychologiques* lorsqu'ils sont au conditionnel présent dans le corpus de presse investi pour cette étude. *Pouvoir* préfère légèrement les verbes de *communication* et *vouloir* les autres. Ils ont cependant beaucoup de verbes en commun si on regarde les indices de spécificité les plus hauts des verbes qui composent ces catégories et qui apparaissent en cooccurrence avec *pouvoir* et *vouloir* :

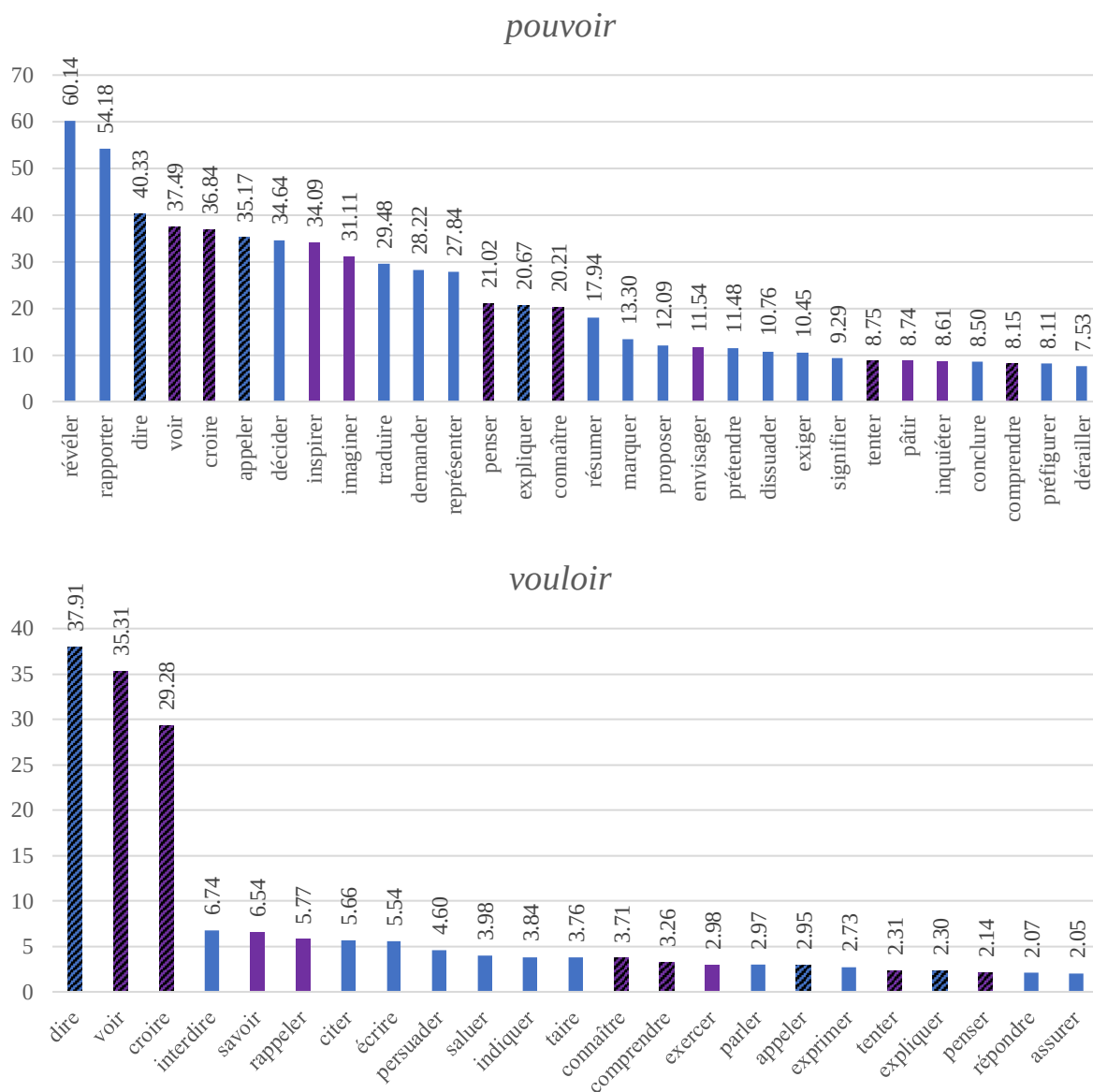


Figure 59 : Les (30) verbes de type communication et psychologique en position de complémentation verbale avec le plus grand indice de spécificité pour pouvoir et vouloir au conditionnel présent dans LM.

Ce sont donc ces cooccurents en commun que nous allons observer de plus près en co(n)texte afin de voir quels types d’usages ressortent pour ces deux modaux qui s’associent de manière similaire à ces deux catégories sémantiques.

Sans grande surprise, *pouvoir* au conditionnel exprime une *possibilité* et donc une valeur modale épistémique :

273. Un lecteur peu informé ***pourrait croire*** qu’en 1975 s’est abattue sur le Cambodge une regrettable catastrophe, sous le nom de « Khmers rouges », et que cette rébellion, venue on ne sait d’où, aurait tué le quart de la population. (LM, 10.08.2010)

274. Mettra-t-il d'accord les deux tutelles ? Une réunion conjointe, prévue le 30 septembre, **pourrait le dire**. (LM, 21.09.2010)
275. Un œil peu aguerri **pourrait** aussi **croire** que des bâtiments publics en ruine ou des écoles détruites sur lesquelles ont été installées des tentes Unicef en guise de classes, ont été atteints par les flots. (LM, 19.08.2010)

Cependant, comme pour les exemples que nous avons relevés au présent pour *pouvoir* et les verbes *psychologiques*, le pronom *on* apparaît également fréquemment au conditionnel présent :

276. On **pourrait croire** que les grèves chez Honda, Hyundai et Foxconn indiquent que la Chine est arrivée au « point de retournement » de Lewis. (LM, 17.06.2010)
277. Véritables laboratoires de création à une période où le cinéma indépendant américain est à bout de souffle, où le cinéma français ne se porte guère mieux, ces nouvelles venues ont comblé un vide chez les spectateurs, suscitant une nouvelle forme de cinéphilie qu'on **pourrait appeler**, pour reprendre l'expression du critique de séries Olivier Joyard, « sériophilie » (LM, 20.05.2010)
278. Pour résumer, on **pourrait dire** que M. Sarkozy dispose de beaucoup de pouvoir quand M. Berlusconi compense le « peu » qu'il a par d'énormes moyens (financiers), grâce auxquels il monnaie ses appuis, ses réseaux, ses vassaux. (LM, 29.12.2010)

Ces exemples montrent un usage de *pouvoir* où le modal, par le conditionnel présent, amène une *suggestion* générale sur une manière de *penser* ou de *dire* quelque chose, il attire à nouveau l'attention sur un fait notable (Sitri 2013) mais de manière plus distancée encore qu'au présent de l'indicatif.

Vouloir au conditionnel avec les verbes de *communication* et *psychologiques* exprime son NS avec une certaine distanciation :

279. Moi, je **voudrais** ne plus **voir** d'enfants exploités dans les rues de Paris. (LM, 15.10.2010)
280. Et franchement, je **voudrais appeler** l'attention sur cette réalité que tout ce qui a été fait depuis 2007 peut être lu avec cette grille. (LM, 16.01.2010)

Dans ces exemples, nous pouvons retrouver le NS d'*intention* voire celui de *volonté*, car il s'agit dans ces cas d'expression d'un *souhait* adouci avec la forme de *vouloir* au conditionnel. Avec *dire*, contrairement aux formes au présent de l'indicatif, nous retrouvons moins souvent le sens de *signifier* ou la fonction de *reformulation* :

281. Je **voudrais dire** un mot de toutes ces questions d'environnement, car là aussi, cela commence à bien faire. (LM, 25.03.2010)
282. Je meurs d'envie qu'on ait une bonne causerie un jour parce qu'il y a des choses que je **voudrais te dire** (LM, 05.11.2010)
283. Aux hommes, aux femmes que cette angoisse-là étirent, on **voudrait dire** que notre nullité n'est pas si complète qu'elle le paraît (LM, 09.07.2010)

A nouveau, l'association ici ne revête pas le sens de la grammaticalisation car elle est précédée par un sujet *animé*. Lorsque ce n'est pas le cas, nous pouvons déceler le sens de *signifier* ou de *reformulation* comme au présent :

284. Tout cela est très dangereux. Si Woerth devait démissionner, cela **voudrait dire** que la rumeur l'a emporté (LM, 15.07.2010)
285. Ça **voudrait dire** que ce serait la responsabilité des Verts si ça ne marchait pas ? (LM, 22.09.2010)

Toutefois, contrairement aux cas où *vouloir* est conjugué au présent de l'indicatif, la locution se grammaticalise dans moins de 50% des cas au conditionnel.

Pour résumer, les usages de *pouvoir* et *vouloir* au conditionnel avec les verbes de *communication* ou *psychologiques* peuvent signaler des emplois en marge des catégories traditionnelles de la modalité. *Pouvoir* permet – comme au présent de l'indicatif – de mettre en avant un fait notable, plus précisément une *parole* ou une *pensée* et *vouloir* se désémantise, encore une fois, lorsqu'il est suivi du verbe *dire*. Cependant, les valeurs modales épistémiques et bouliques sont ici plus souvent en vigueur que pour les formes au présent de l'indicatif relevées pour ces deux verbes dans la section consacrée à ce temps verbal dans LM.

7.2.3.10. Les complémentations verbales de type verbes de entrée/sortie, don/privation et communication avec devoir au conditionnel présent

Devoir semble être partagé entre plusieurs catégories de verbes lorsqu'il est conjugué au conditionnel présent. Ainsi, tant des verbes d'*entrée/sortie*, de *don/privation* ou encore de *communication* s'associent de manière importante avec *devoir* :

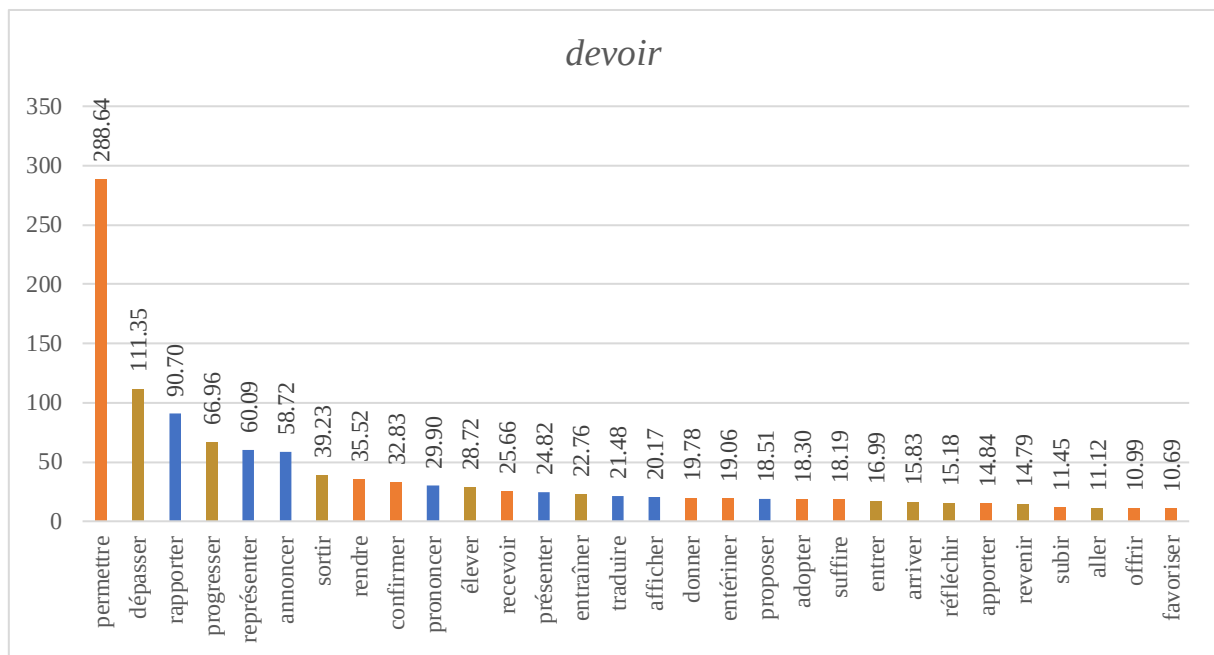


Figure 60 : Les 30 verbes de type entrée/sortie, don/privation et communication en position de complémentation verbale avec le plus grand indice de spécificité pour *devoir* au conditionnel présent dans LM.

Le verbe *permettre* ressort de manière particulière avec un indice de spécificité de 288,64. Les usages de *devoir* avec ce verbe sont de l'ordre de la valeur modale épistémique :

286. Ses projets **devraient** lui **permettre** de fabriquer un million de véhicules par an d'ici à 2015. (LM, 01.10.2010)
287. Celui-ci a donc donné son feu vert : les conditions météorologiques **devraient** lui **permettre** d'atteindre le pôle Nord en trois jours. (LM, 06.04.2010)
288. Les travaux **devraient permettre** de doubler le nombre de visiteurs susceptibles d'être accueillis en même temps - 20 000 dont 10 000 sous la seule nef, au lieu de 5 800 aujourd'hui. (LM, 22.04.2010)

Cette valeur modale épistémique s'interprète non seulement à l'aide du mode du conditionnel, qui, comme nous l'avons déjà soulevé à plusieurs reprises confère à *devoir* cette lecture sémantique, mais aussi grâce au verbe *permettre* qui suit. En effet, celui-ci, dans son interprétation la plus sommaire, se rapproche du NS de *possibilité* de *pouvoir* : on donne la *possibilité* de faire quelque chose. Ici, le cotexte semble influencer la lecture sémantique de *devoir*.

Cependant, même lorsque *devoir* est suivi d'un autre verbe se trouvant dans la figure 61, nous pouvons toujours l'interpréter comme véhiculant une valeur modale épistémique :

289. Didier Fontenille pense donc que la situation actuelle ne **devrait** pas **donner** lieu à une épidémie - trente cas minimum. (LM, 28.09.2010)
290. Le festival accueillera aussi un projet original du producteur Daniel Lanois, Black Dub et **devrait rendre** un hommage à Django Reinhardt. (LM, 15.03.2010)
291. A la RATP, on annonce un trafic normal ou quasi normal, sauf sur les RER A et B. Le trafic aérien **devrait subir** des perturbations (50 % des vols seront annulés à Orly, 30 % ailleurs). (LM, 20.10.2010)

Les effets de sens sont ici pour ces trois exemples de l'ordre de la *supposition* ou de la *prédiction*.

C'est d'ailleurs ces *supposition* et *prédiction* qui se retrouvent fréquemment dans les usages de *devoir* avec les verbes de communication :

292. Si nous devons entrer en contact avec les extraterrestres, qui **devrait représenter** l'humanité ? (LM, 30.09.2010)
293. Mme Aubry **devrait annoncer** si elle est candidate à l'élection présidentielle. (LM, 27.08.2010)
294. Le tribunal administratif **devrait se prononcer** avant les vacances d'été. (LM, 14.06.2010)
295. Sony **devrait** également **présenter** officiellement sa manette à détection de mouvements, le Motion controller. (LM, 08.05.2010)
296. La réforme en cours prévoit notamment l'externalisation du traitement des magazines. Elle **devrait se traduire** par un plan de départs volontaires - le nombre de départs attendus par la direction n'est pas précisé. (LM, 15.12.2010)
297. Aucun d'entre eux n'a, pour l'instant, signé d'accord avec Apple, mais Numilog (filiale numérique d'Hachette) **devrait proposer** courant mai 10 000 titres des maisons appartenant au groupe (Grasset, Fayard, Stock, Lattès...) et d'Albin Michel, qui seront ainsi téléchargeables sur iPhone ou iPad. (LM, 16.04.2010)

Nous avons ici illustré de nombreux exemples pour montrer que contrairement aux autres verbes modaux que nous étudions dans cette thèse, *devoir* ne semble pas avoir d'usages qui se distinguent des emplois traditionnels qu'on lui attribue alors que *pouvoir*, *falloir* et *vouloir*, avec les verbes de *communication* et *psychologiques* peuvent montrer des phénomènes de grammaticalisation.

En résumé, il semble que l'utilisation du conditionnel pour *devoir* dans LM soit limitée à une valeur modale épistémique, quelle que soit la catégorie sémantique du verbe à l'infinitif qui est modalisé, du moins d'après les échantillons que nous avons examinés. Cette tendance ne se

retrouve pas avec les autres modaux, qui peuvent prendre des lectures sémantiques moins conventionnelles lorsqu'ils sont associés à des verbes de *communication* et *psychologiques*.

7.2.3.11. Les complémentations verbales de type verbes psychologiques avec falloir au conditionnel présent

Falloir au conditionnel attire majoritairement des verbes *psychologiques*, comme c'était déjà le cas pour les formes lemmatisées et pour les formes du présent de l'indicatif :

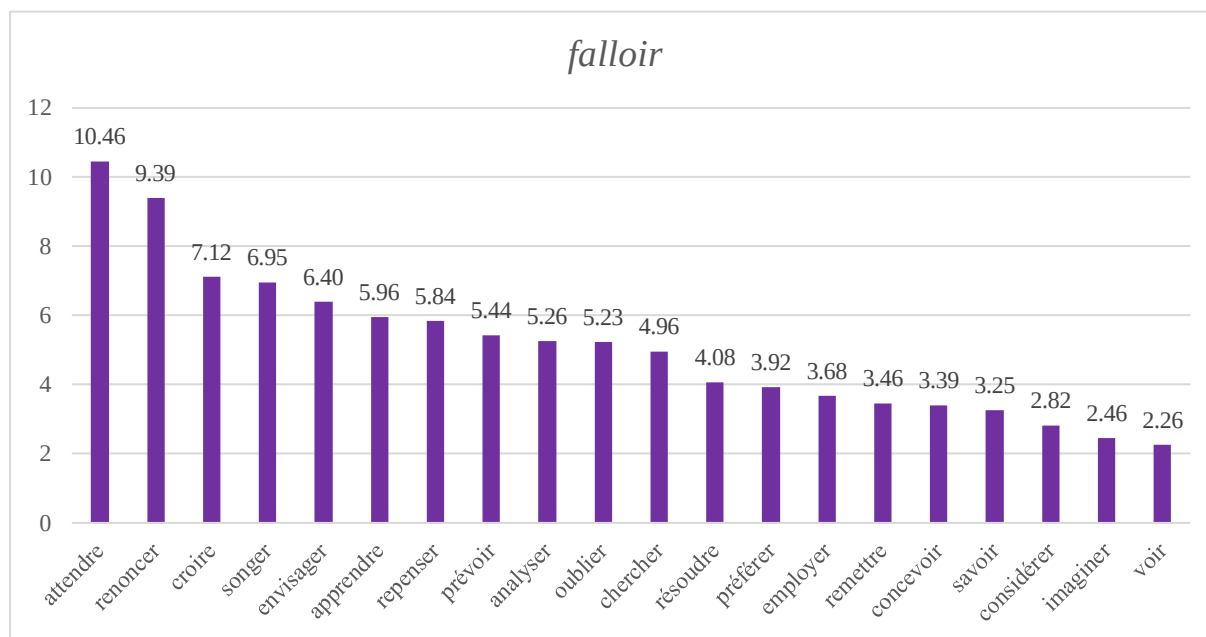


Figure 61 : Les verbes de type psychologique en position de complémentation verbale avec le plus grand indice de spécificité pour falloir au conditionnel présent dans LM.

Falloir au conditionnel avec les verbes *psychologiques* agit comme un marqueur qui permet de présenter une pensée alternative : avec ce que nous venons d'apprendre, ou ce que nous savons déjà, devrions-nous pas avoir une autre pensée ?

298. « Liberté, Égalité, Fraternité » : la devise de la République correspond-elle encore à la réalité ? Ou bien **faudrait**-il se **résoudre** à en supprimer le deuxième terme, tant il apparaît que nous vivons, chaque jour davantage, dans une société inégalitaire ? (LM, 11.03.2010)
299. Mais pour commencer, au lieu de singulariser l'une des sources d'énergie, c'est l'ensemble du mix énergétique qu'il **faudrait considérer**. (LM, 23.08.2010)
300. A ce compte, il va devenir difficile de discerner ce qui n'est pas un roman. Il **faudrait** donc **prévoir** de changer bientôt d'étagère Einstein, Darwin, le Pentateuque et tutti quanti. (LM, 27.08.2010)
301. A l'en croire, il ne **faudrait** y **voir** qu'un « signe privé », comparable à l'engagement chrétien de l'abbé Pierre. (LM, 06.02.2010)

En ceci, il semble qu'il véhicule une valeur modale épistémique puisqu'il présente une *possibilité* de penser autrement¹⁸⁴.

Comme au présent, il peut également fonctionner comme un marqueur de focalisation¹⁸⁵ mais de façon plus distanciée :

302. Un Facebook à la française, le site Copains d'avant ? Et bien oui. Mais il ne **faudrait** pas l'**oublier** : la plate-forme française a vu le jour bien avant que l'idée n'ait germé dans l'esprit de Mark Zuckerberg. (LM, 22.11.2010)
303. Avant de construire ce château espagnol en Afrique du Sud, il **faudrait songer** à ne pas perdre contre l'Uruguay. Ce qui est loin d'être gagné. (LM, 05.03.2010)

En somme, *falloir* au conditionnel, suivi par des verbes *psychologiques*, permet de présenter des pensées alternatives ou de ponctuer le discours, de manière plus adoucie qu'au présent de l'indicatif. Ces recherches et ces exemples permettent d'étayer ce qui a été dit et confirmé déjà à son sujet, c'est-à-dire que ce modal a tendance à se grammaticaliser lorsqu'il est suivi de verbes de ce type, et ce, même au conditionnel.

7.2.4. Les résultats dans WIKI

Comme pour l'étude des sujets nominaux, nous allons comparer les résultats obtenus dans le corpus de presse à un corpus encyclopédique, en l'occurrence Wikipédia. Ces résultats permettront de souligner si les conclusions que nous avons tirées au sujet des verbes modaux selon leur complémentation verbale peuvent être généralisées ou du moins étayées à l'aide d'un autre texte. Il sera notamment intéressant de voir si les verbes de *communication* et verbes *psychologiques* confèrent à *pouvoir*, *falloir* et *vouloir* des usages aussi éloignés de leurs NS respectifs que dans le corpus de presse.

¹⁸⁴ Nous pouvons rapprocher ce type d'effet de sens de la notion de *désirabilité*. Selon Nuyts (2016), « [l]a modalité déontique peut être définie comme une indication du degré de désirabilité morale de l'état de choses exprimé dans l'énoncé, typiquement mais pas nécessairement au nom de la locutrice (les locutrices peuvent rapporter les jugements déontiques d'autres personnes) » (Nuyts 2016 : 36). Dans ce type d'exemples, la locutrice montre une façon plus *désirable* de penser quelque chose.

¹⁸⁵ Nous utilisons la notion *marqueur de focalisation* car ici *falloir* ne semble pas agir comme un MD selon les critères que nous avons relevés en 4.1.1. même s'il semble être en voie de grammaticalisation dans ce type de construction.

7.2.4.1. Les répartitions des verbes en cooccurrences significatives pour les formes lemmatisées

Les profils combinatoires des verbes modaux avec leurs complémentations verbales sont différents dans le corpus WIKI par rapport à ce que nous avons déjà observé pour le corpus LM :

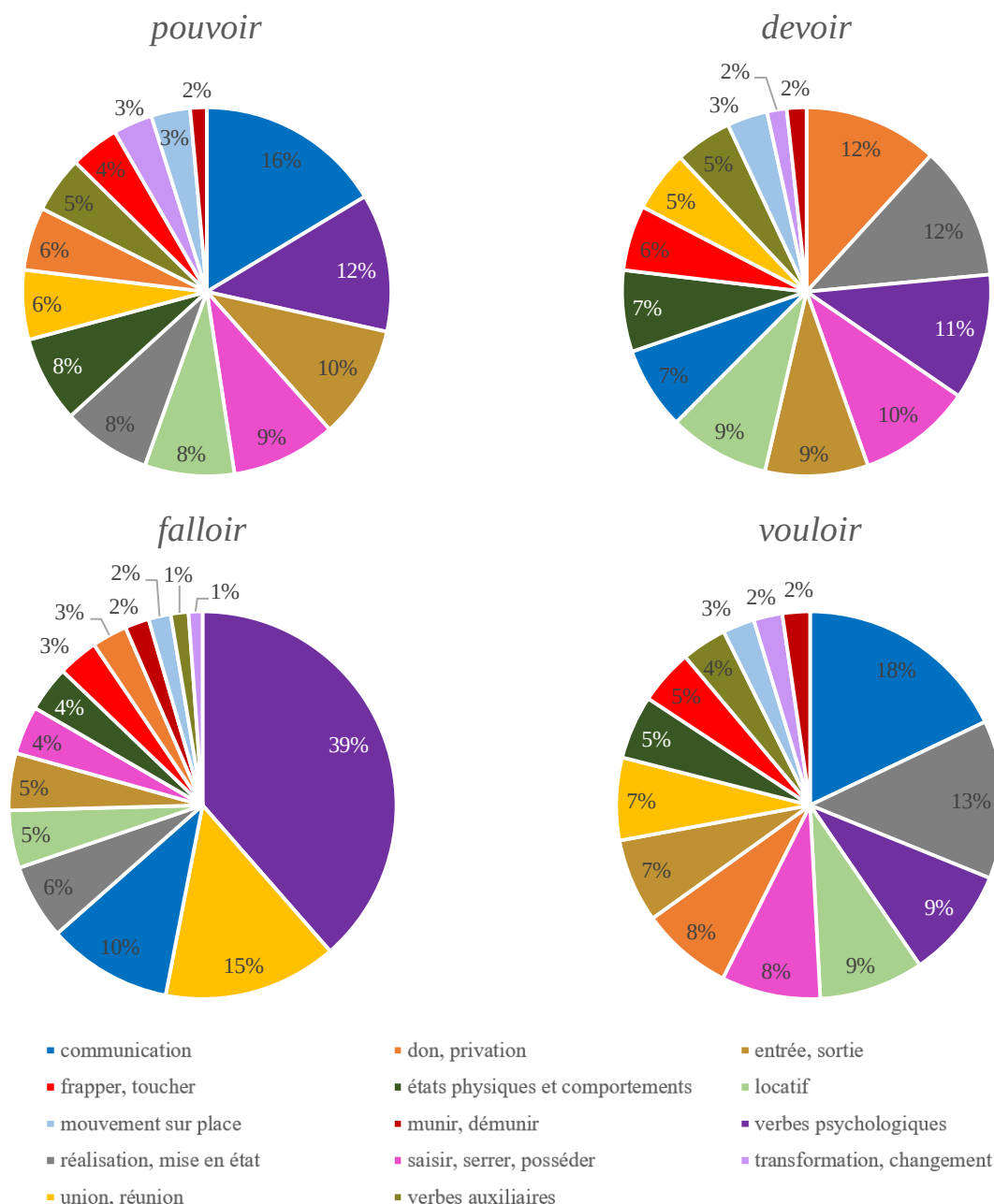


Figure 62 : Répartition des verbes à l'infinitif en position de complémentation verbale selon Dubois et Dubois-Charlier (1997) pour les formes lemmatisées des modaux dans le corpus WIKI.

Pouvoir attire à nouveau majoritairement des verbes de *communication* et *psychologiques* mais avec cette fois-ci une légère préférence pour les verbes de *communication* par rapport au corpus LM. *Devoir*, quant à lui, préfère les verbes de *don/privation* ainsi que les verbes de

réalisation/mise en état avant les verbes *psychologiques*. Les verbes de *communication* arrivent encore plus loin dans le classement. *Falloir* attire toujours majoritairement des verbes *psychologiques* mais cette fois-ci de manière encore plus marquée vu que 39% de ses cooccurrences significatives concernent des verbes de ce type. Les verbes d'*union/réunion* arrivent en deuxième position avant les verbes de *communication*. Quant à *vouloir*, les verbes de *communication* prédominent une fois de plus parmi les cooccurents significatifs, suivis par les verbes de *réalisation/mise en état* puis les verbes *psychologiques*.

Nous allons analyser des occurrences des quatre modaux séparément au vu de leurs profils combinatoires distincts : pour *pouvoir*, nous allons étudier ses combinaisons avec les verbes de *communication* et les verbes *psychologiques*, pour *devoir*, également des occurrences avec les deux catégories qui arrivent en tête, c'est-à-dire les verbes de *don/privation* ainsi que ceux de la catégorie *réalisation/mise en état*. Pour *falloir*, ce sont les usages avec les verbes *psychologiques* qui seront décortiqués. Pour *vouloir*, nous allons nous focaliser sur les verbes de *communication* essentiellement.

7.2.4.2. Les complémentations verbales de type verbes de communication et psychologiques avec pouvoir pour ses formes lemmatisées

Les indices de spécificité que *pouvoir* présente avec les verbes de *communication* et *psychologiques* sont relativement élevés comparés à d'autres indices que nous avons présentés tout au long de ce travail :

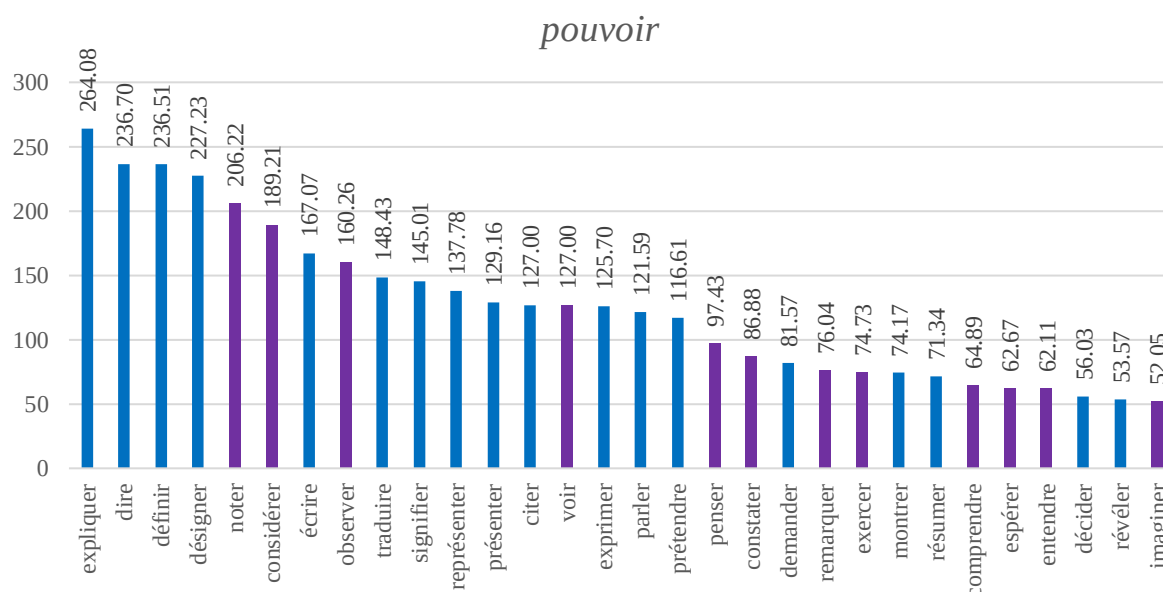


Figure 63 : Les 30 verbes de type communication et psychologiques en position de complémentation verbale avec le plus grand indice de spécificité pour pouvoir pour ses formes lemmatisées dans le corpus WIKI.

Nous retrouvons des usages classiques de *pouvoir* comme la *capacité*, générale en 304, d'une personne en 305 ainsi que d'un événement *capable* de se produire (dans le sens qu'il est *possible* que cela arrive) en 306 :

- 304. On **peut voir** au Musée Ingres une vieille thériaque où figure l'inscription « faz » (pour « fait »), deux cheminées où figure le rébus de Cahors illustré en occitan par un chien (« can ») et un ours (« ors »), la vieille cloche « La Berlòca », quelques pierres tombales, le cadran solaire situé sur l'église Saint-Jacques avec son inscription « tard o d'ora vendrà l'ora ». (WIKI)
- 305. Lorsque William Grindal devint son tuteur en 1544, Élisabeth **pouvait écrire** en anglais, en latin et en italien et, sous son enseignement, elle progressa en français et en grec. (WIKI)
- 306. [...] trois ans auparavant, Tchernobyl avait rappelé à l'ensemble de la planète que l'énergie nucléaire de fission **pouvait présenter** des dangers considérables et les programmes électro-nucléaires engagés au lendemain du premier choc pétrolier de 1973 commençaient à être de plus en plus critiqués par l'opinion publique [...] (WIKI)

Cependant, avec les verbes de *communication* ou verbes *psychologiques*, nous retrouvons, dans la plupart des cas, des énoncés où *pouvoir* se grammaticalise, du moins, dans l'échantillon étudié. Ainsi, la construction qu'il compose sert à attirer l'attention de la lectrice ou de l'interlocutrice sur un fait notable (Sitri 2013) :

- 307. De nombreux grands espaces ouverts sont également présents dans le centre-ville d'Amsterdam, au premier rang desquels on **peut citer** le Dam, grande place sur laquelle sont situés le palais royal et le « National Monument », ou encore Museumplein, une grande zone recouverte de pelouse où sont regroupés les musées du Rijksmuseum, le musée Van Gogh et le Stedelijk Museum. (WIKI)
- 308. On **peut** aussi **parler** du travail de Jacques Offenbach (auteur des « Contes d'Hoffmann »), compositeur parisien né en Allemagne qui s'imposa comme le maître de l'opéra-comique français du XIX^e siècle, appelé opéra-bouffe. (WIKI)
- 309. Le fameux Discours de la Méthode, écrit en français, chercha à décrire une manière déductive de traiter les problèmes, beaucoup moins fondée sur l'intuition. On **peut dire** qu'il marque le début de la démarche des sciences dites « exactes », fondées sur les raisonnements logiques de déduction. (WIKI)
- 310. Sur cette mixtape, on **peut remarquer** la collaboration de plusieurs groupes tels que Chiens de paille, Psy 4 De La Rime ou encore MC Arabica. (WIKI)

311. Le mobilier a gardé ici son état du Second Empire : on **peut** par exemple **noter** des feux en bronze patiné représentant « Vénus et Adonis », sur un modèle de la Renaissance (acquis en 1860). (WIKI)
312. Parmi les divinités mineures figurent tous ceux que nous **pouvons considérer** comme des démons ou des génies. (WIKI)

Dans la grande majorité des cas, il est précédé par le pronom *on* et plus rarement *nous*. Ici, le NS de *possibilité* est difficilement récupérable : *pouvoir* sert simplement ici comme marqueur d'attention tout en mettant de la distance sur le fait relevé.

Pouvoir sert également de marqueur d'alternative :

313. Comme en français, les albanais prononcent parfois la suite « en » comme [ã] (a nasal) à la place de [ɛn], comme dans « vend » (lieu, endroit), qui **peut se dire** soit /vɛnd/, soit /vãd/. Cela est aussi valable avec « ën », tënd (ton, votre) pouvant se prononcer /tənd/ ou /tãd/. (WIKI)
314. La quasi-totalité des noms féminins ont pour terminaison le son [a], qui **peut s'écrire** а ou я, et les noms masculins une consonne, y compris la plupart des noms personnels. (WIKI)
315. L'accolement de deux noms **peut désigner** le résultat de l'addition (comme dans « dix-sept ») ou de la multiplication (comme dans « quatre-vingts ») des entiers correspondants. (WIKI)

Dès lors, *pouvoir* sert dans ce type d'usage la plupart du temps comme marqueur d'attention avec les verbes de *communication* et *psychologiques* dans le corpus WIKI même si, une fois de plus, ce modal est hautement polysémique et peut également véhiculer des effets de sens beaucoup plus usuels et issus de des catégories modales traditionnelles.

7.2.4.3. Les complémentations verbales de type verbes de don/privation et réalisation/mise en état avec devoir pour ses formes lemmatisées

Comme pour *devoir* au conditionnel dans LM, c'est le verbe *permettre* qui apparaît en tête de liste lorsque nous observons la catégorie des verbes de *don/privation* et celle de *réalisation/mise en état* pour les formes lemmatisées dans WIKI :

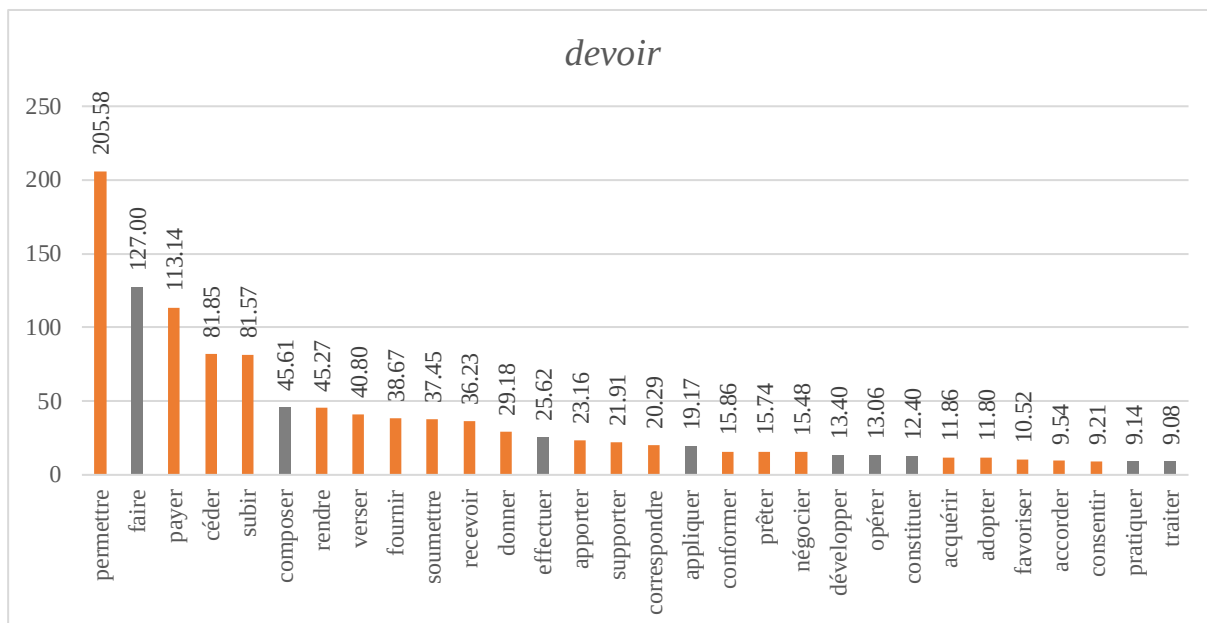


Figure 64 : Les 30 verbes de type don/privation et réalisation/mise en état en position de complémentation verbale avec le plus grand indice de spécificité pour *devoir* pour ses formes lemmatisées dans le corpus WIKI.

Le verbe *permettre* semble exercer une influence sur la lecture sémantique de *devoir* car, dans les exemples qui suivent, nous pouvons comprendre *devoir* comme interprétant une *prédiction* et véhiculant ainsi une valeur modale épistémique :

316. Le 1^{er} février 2006, l'Argentine et le Brésil signent, après près de trois ans de négociations, un accord qui **doit permettre** de protéger les secteurs de production qui pourraient être trop durement affectés par la compétition du pays voisin. (WIKI)
317. Les techniciens de Peenemünde avaient établi des plans de ces énormes installations qui **devaient permettre** d'abriter sous plusieurs mètres de béton le carburant et le comburant, des locaux techniques permettant de produire l'oxygène liquide et d'effectuer des tests, des entrepôts pour les missiles, un système de transport permettant de véhiculer les fusées, et des baraquements permettant d'héberger 250 à 300 spécialistes, le tout protégé par des batteries anti-aériennes. (WIKI)

Cependant, c'est uniquement avec le verbe *permettre* parmi ceux de la figure 65 que *devoir* semble acquérir cet effet de sens. En effet, en cooccurrence avec d'autres verbes, c'est la valeur modale déontique qui ressort :

318. Athéna convainc cette dernière d'allaiter le bébé, mais Héraclès tête trop goulûment et Athéna **doit** le **rendre** à sa mère. (WIKI)
319. Pour pouvoir graver un CD-RW, les vitesses de ce CD-RW **doivent correspondre** à celles du graveur. (WIKI)

320. De fait, ce n'est pas dans le cadre de la relativité restreinte que l'on **doit appliquer** la loi de Hubble-Lemaître, mais celui de la relativité générale. (WIKI)
321. En 1916, Venustiano Carranza était le seul à pouvoir prétendre au pouvoir suprême, même s'il ne contrôlait pas l'ensemble du pays et qu'il **devait faire** face à d'énormes problèmes socio-économiques. (WIKI)

Ainsi, *devoir*, une fois de plus, se rattache fortement à son NS de *nécessité*. Il semble toutefois posséder la particularité de dénoter un effet de sens particulier lorsqu'il modalise le verbe *permettre*, soit la *prédiction*, et ce, quel que soit le corpus et le temps ou le mode verbal auquel il est conjugué.

7.2.4.4. Les complémentations verbales de type verbes psychologiques avec falloir pour ses formes lemmatisées

Les verbes *psychologiques* représentent 39% des verbes spécifiques en position de complémentation verbale pour *falloir* en tant que lemme dans WIKI.

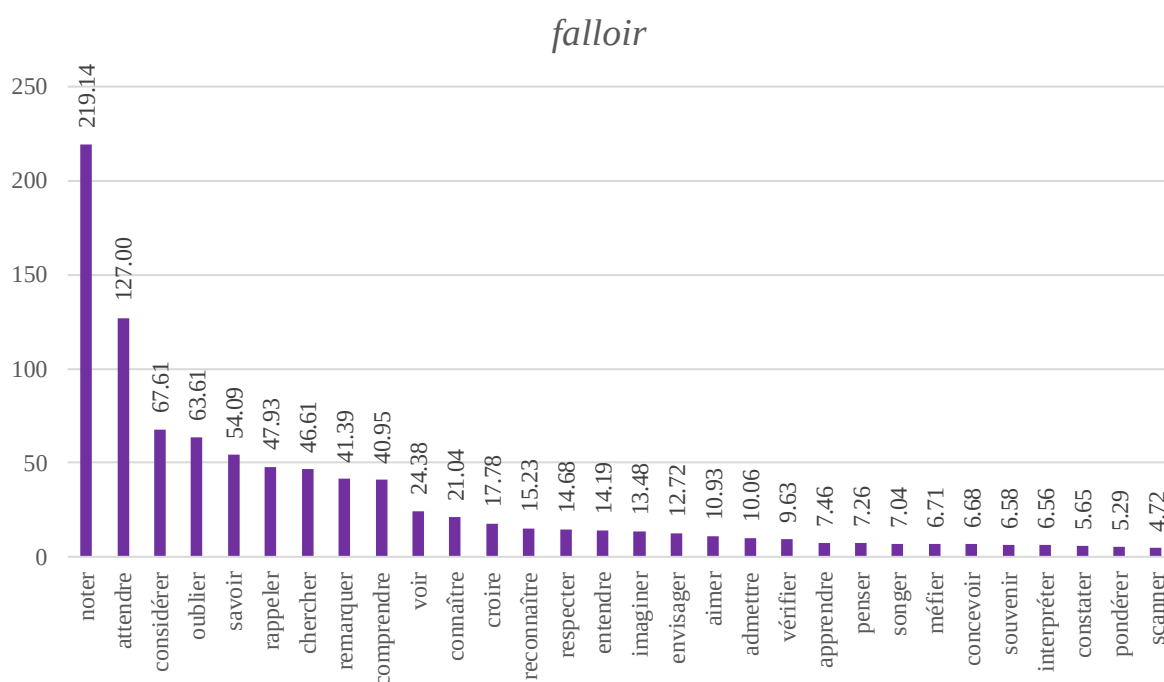


Figure 65 : Les 30 verbes de type psychologique en position de complémentation verbale avec le plus grand indice de spécificité pour *falloir* pour ses formes lemmatisées dans le corpus WIKI.

Le verbe possédant l'indice le plus important dans ce corpus avec *falloir* en tant que lemme, est *noter*. Dans ces cas, une nouvelle fois, *falloir* semble agir comme un marqueur de focalisation, au même titre qu'avec les autres associations que nous avons citées dans les autres sections :

322. Il **faut** également **noter** qu'avec l'accroissement du nombre de locuteurs, l'espéranto est devenu la langue maternelle d'enfants issus de couples espérantophones. (WIKI)
323. Il **faut noter** que la plupart des mots se terminant par « e » sont prononcés avec un coup de glotte à la fin et se terminent donc par une syllabe fermée. (WIKI)
324. Il **faut noter** d'ailleurs que ses premiers romans sont situés dans des mondes d'hommes, exprimant un point de vue masculin. (WIKI)

Dans ces exemples, *falloir* met en évidence un fait général. Le NS de *contrainte* pourrait être récupéré si l'on admet que dans ces cas, *falloir* permet de signifier que « nous sommes sous la contrainte de connaître un fait afin d'en comprendre d'autres ». Cette inférence n'est cependant pas évidente et c'est pourquoi nous pouvons considérer ici ces constructions comme relevant de la grammaticalisation, dans une phase où le NS n'est plus percevable.

Le même type d'interprétation convient pour le deuxième verbe ayant l'indice le plus grand en position de complémentation verbale, soit *attendre* :

325. Après l'occupation allemande et le bombardement du dépôt de tramways, les bus prennent la succession, et il **faut attendre** 1962 pour que le réseau dispose de plusieurs lignes. (WIKI)
326. Les paysages occupent une part primordiale dans l'art en Extrême-Orient, notamment en Chine et au Japon, mais il **faudra attendre** la Renaissance en Europe pour voir les paysages prendre de l'importance dans la peinture. (WIKI)
327. Il **a fallu attendre** le début du XX^e siècle pour que le principe de bivalence soit clairement remis en question de plusieurs façons différentes. (WIKI)

Dans ces occurrences, est-il réellement *nécessaire* d'*attendre* ? Nous pouvons interpréter le verbe *falloir* comme signifiant qu'« avant un événement ou une date donnée, certaines pensées ou faits n'étaient pas envisageables », et qu'il est donc « impératif d'attendre cet événement spécifique pour que cela devienne possible ». Cependant, cette interprétation n'est pas totalement assurée. Il semble plutôt que l'expression *falloir attendre* fonctionne simplement comme un marqueur textuel servant à indiquer un fait dans le temps. Avec ce type d'interprétation, *falloir* semble à nouveau se grammaticaliser et agit comme un marqueur de focalisation sur un certain fait temporel.

Avec les autres verbes *psychologiques*, *falloir* fonctionne, comme soulevé précédemment, également comme marqueur de focalisation :

328. Si les cotes peuvent grimper, il ne **faut** pas **oublier** que le prix peut varier, pour les archets anciens, dans des proportions assez importantes, selon l'état et les éventuelles restaurations subies, ainsi que selon les qualités du bois utilisé par le fabricant. (WIKI)
329. Or, avant de pouvoir étudier des institutions sociales, il **faut savoir** en quoi elles consistent exactement. (WIKI)

En somme, *falloir*, comme nous l'avions déjà relevé dans le corpus LM, a tendance à se transformer en marqueur de focalisation avec les verbes *psychologiques*. Il permet d'attirer l'attention de la lectrice sur un fait particulier.

7.2.4.5. Les complémentations verbales de type verbes de communication avec vouloir pour ses formes lemmatisées

Sans surprise, c'est le verbe *dire* qui arrive en tête des verbes de *communication* ayant le plus grand indice de spécificité en position de complémentation verbale pour *vouloir* :

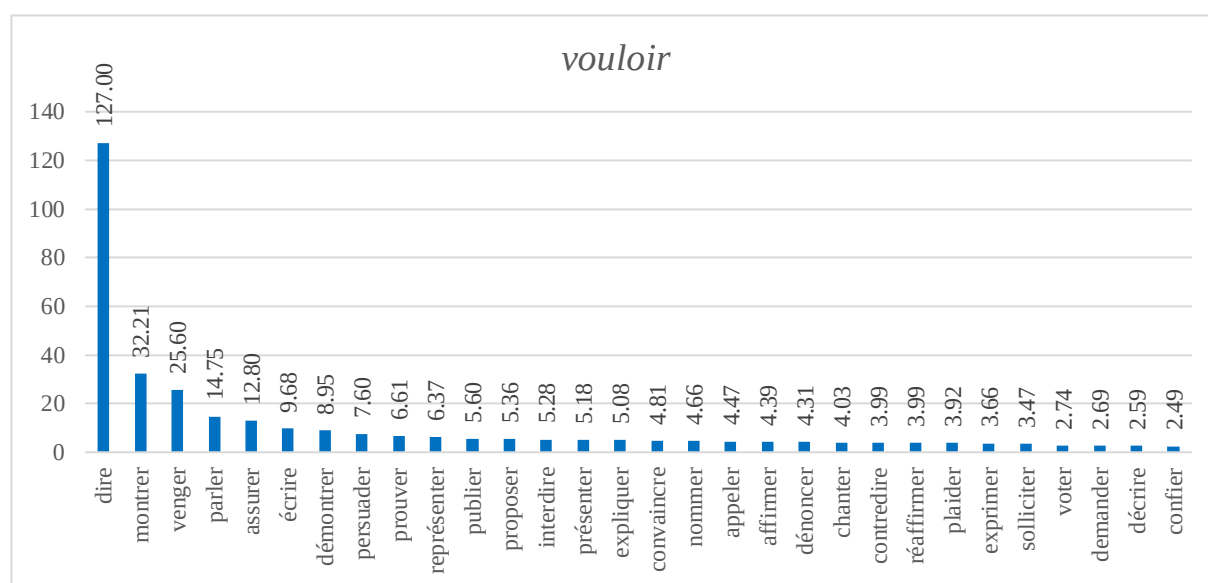


Figure 66 : Les 30 verbes de communication en position de complémentation verbale avec le plus grand indice de spécificité pour *vouloir* pour ses formes lemmatisées dans le corpus WIKI.

On retrouve bien sûr la locution *vouloir dire* dans la majorité des occurrences :

330. Si l'ensemble des commentateurs s'accordent sur quelques traits généraux, les discussions de détail ne permettent pas de décider de manière certaine de ce que **voulait dire** Anaximandre. (WIKI)
331. Ça ne **veut pas dire** qu'ils ne peuvent pas changer, mais il faut un effort laborieux pour le faire. (WIKI)
332. À l'origine, « gouyé » **veut dire** « petite mare ». (WIKI)

Dans ces cas-là, nous retrouvons donc le sens de *signifier* ou de reformulation lorsque ces deux verbes cooccurrent.

Nous avons pu vérifier toutes les occurrences de *vouloir* avec *dire* et dans seulement moins de 3% des cas, la cooccurrence n'est pas grammaticalisée :

- 333. Je **veux dire** ici solennellement que si des communautés ont été prises pour cibles à Montauban comme à Toulouse, ce sont des enfants, des soldats des Français qui ont été assassinés. (WIKI)
- 334. Les personnes qui présentent une aphasie de Broca savent ce qu'elles **veulent dire**, mais ont de grandes difficultés à parler. (WIKI)

Enfin, comme dans les autres exemples que nous avons observés dans LM, lorsque *vouloir* cooccur avec d'autres verbes de cette catégorie, nous retrouvons le NS d'*intention/volonté* aisément :

- 335. Stanley Milgram s'oppose fortement aux interprétations qui **voudraient expliquer** les résultats expérimentaux par l'agressivité interne des sujets. (WIKI)
- 336. Enfin, Descartes **veut expliquer** tous les phénomènes de la nature : il étudie les êtres vivants et fait de nombreuses dissections à Amsterdam pendant l'hiver 1631 - 1632. (WIKI)
- 337. [...] tandis que les auteurs du genre imaginent généralement des êtres à l'esprit moyenâgeux possédant une technologie supérieure, Ivan Efremov **a voulu décrire** une société dont la conscience est en adéquation avec son développement technologique. (WIKI)
- 338. Je **veux vraiment parler** finnois mais je n'en suis pas capable. (WIKI)
- 339. Mitterrand, qui reste sur deux échecs en 1974 et 1978, reste très évasif sur son éventuelle participation à la prochaine élection, tandis que Rocard, qui **veut persuader** qu'il est le seul à pouvoir mener les socialistes à la victoire, se montre bien plus déterminé. (WIKI)

Vouloir semble ainsi avoir un comportement stable avec les verbes de *communication* : quel que soit le temps ou le mode verbal auquel il est conjugué et le corpus dans lequel il apparaît, il se grammaticalise avec *dire* et exprime son NS d'*intention* ou de *volonté* avec les autres verbes de cette catégorie.

7.2.4.6. Les répartitions des verbes en cooccurrences significatives pour les formes au présent de l'indicatif

Les profils combinatoires des verbes au présent dans WIKI ressemblent beaucoup à ceux que nous avons relevés pour les formes lemmatisées :

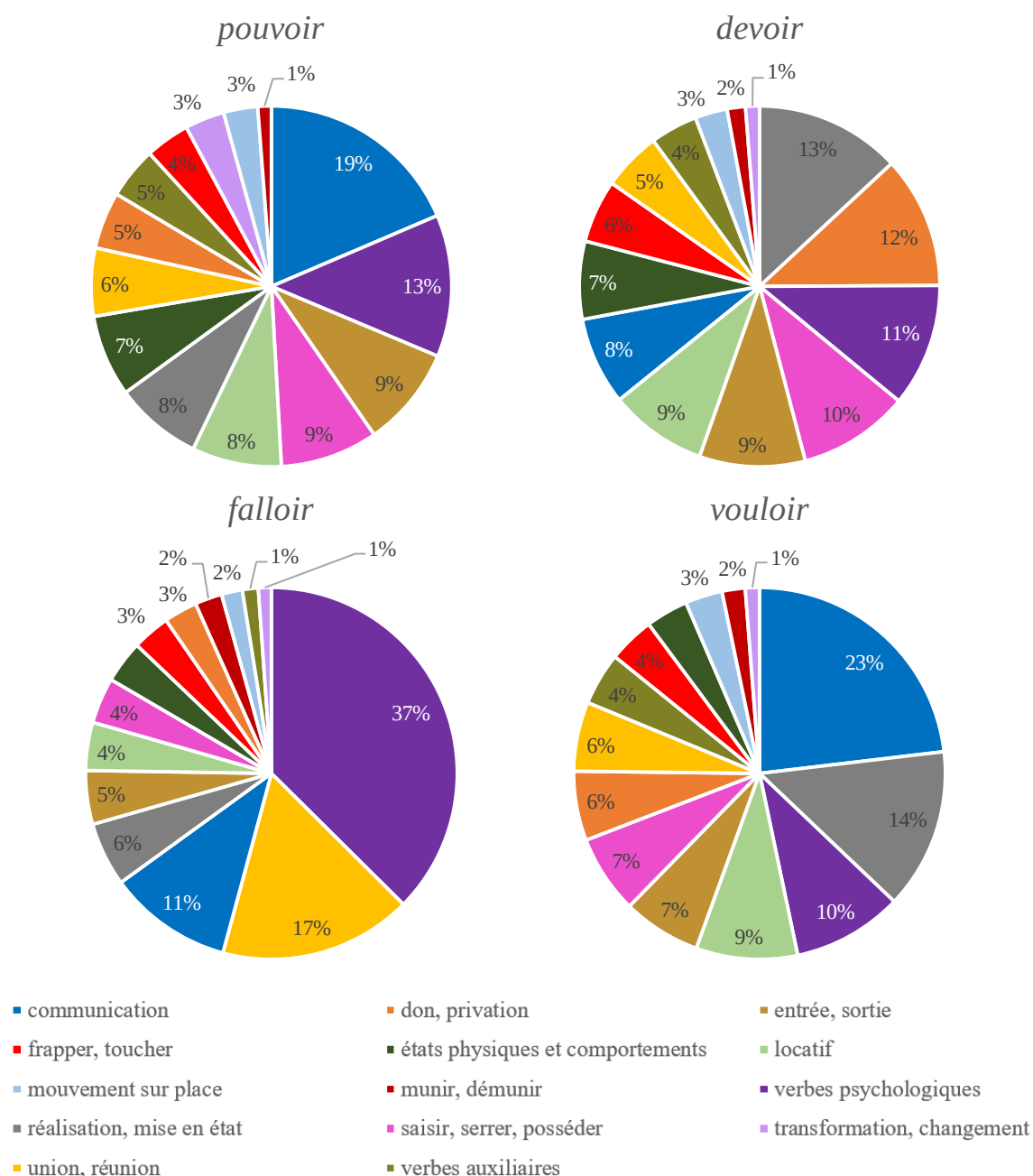


Figure 67 : Répartition des verbes à l'infinitif en position de complémentation verbale selon la classification Dubois et Dubois-Charlier (1997) pour les formes du présent de l'indicatif des modaux dans le corpus WIKI.

Pouvoir, *falloir* et *vouloir* ont les mêmes trois premières catégories représentées de façon majoritaire que lorsque nous avons étudié les profils combinatoires de leurs formes

lemmatisées. *Devoir* inverse ses deux premières catégories en représentant de manière majoritaire ici les verbes de type *réalisation/mise en état* puis ceux de *don/privation*.

Au vu des résultats similaires, et puisque nous avons déjà relevé énormément d'exemples au présent de l'indicatif dans la partie sur les formes lemmatisées, nous ne nous focaliserons ici pas sur les indices de spécificité et les occurrences dans cette section et proposons directement d'entrer dans les profils combinatoires des modaux au conditionnel présent dans WIKI afin de voir si avec ce mode verbal nous pouvons voir émerger d'autres particularités.

7.2.4.7. Les répartitions des verbes en cooccurrences significatives pour les formes au conditionnel présent

Au conditionnel présent dans WIKI, nous avons des profils combinatoires qui diffèrent de ceux que nous avons obtenus pour les lemmes ou les formes au présent de l'indicatif :

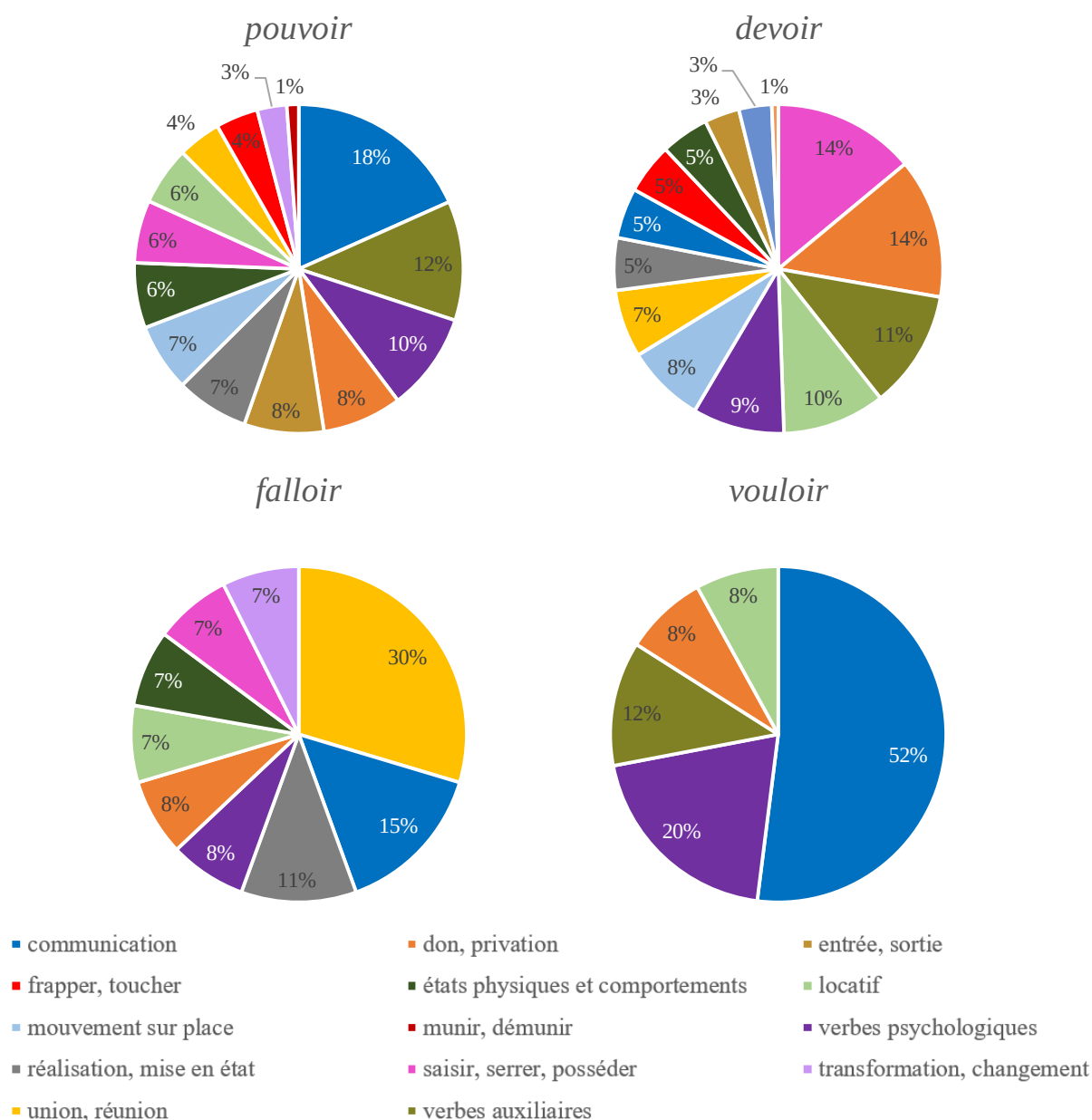


Figure 68 : Répartition des verbes à l'infinitif en position de complément verbale selon la classification de Dubois et Dubois-Charlier (1997) pour les formes du conditionnel présent des modaux dans le corpus WIKI.

Pouvoir attire majoritairement des verbes de *communication* mais cette fois-ci, ce sont les verbes *auxiliaires*¹⁸⁶ qui arrivent en seconde position avant les verbes *psychologiques*. *Devoir*, pour sa part, préfère les verbes de type *saisir/serrer/posséder* puis les verbes de *don/privation* en seconde position et les *auxiliaires* en troisième. *Falloir*, lui, diffère des remarques qu'on a pu soulever à son sujet jusqu'à maintenant puisqu'il préfère dans cette configuration les verbes de type *union/réunion*. Les verbes *psychologiques* ne représentent, dans cette configuration, que 8% des cooccurents spécifiques à *falloir*. En dernier lieu, le verbe *vouloir* est

¹⁸⁶ Rappelons que nous avons omis ici les auxiliaires *être* et *avoir* de ces chiffres afin de ne pas avoir une surreprésentation de cette catégorie.

majoritairement accompagné de verbes de *communication*, avec plus d'un verbe cooccurrent spécifique sur deux en position de complémentation verbale, qui se classe dans cette catégorie. Les verbes *psychologiques* représentent une occurrence spécifique sur cinq.

Ce sont donc ces particularités-là que nous allons étudier : en premier lieu, les occurrences de *pouvoir* avec les verbes de *communication* ; pour *devoir*, celles avec les verbes de type *saisir/serrer/posséder* ainsi qu'avec les verbes de la catégorie *don/privation* ; pour *falloir*, nous nous centrerons sur les cooccurrences avec les verbes représentant la catégorie *union/réunion*. Enfin, la partie consacrée à *vouloir* se centrera sur ses occurrences avec les verbes de *communication*.

7.2.4.8. Les complémentations verbales de type verbes de communication pour pouvoir au conditionnel présent

Comme pour les formes lemmatisées, ce sont les verbes *expliquer* et *dire* qui possèdent le plus grand indice de spécificité lorsque *pouvoir* est usité au conditionnel présent dans WIKI :

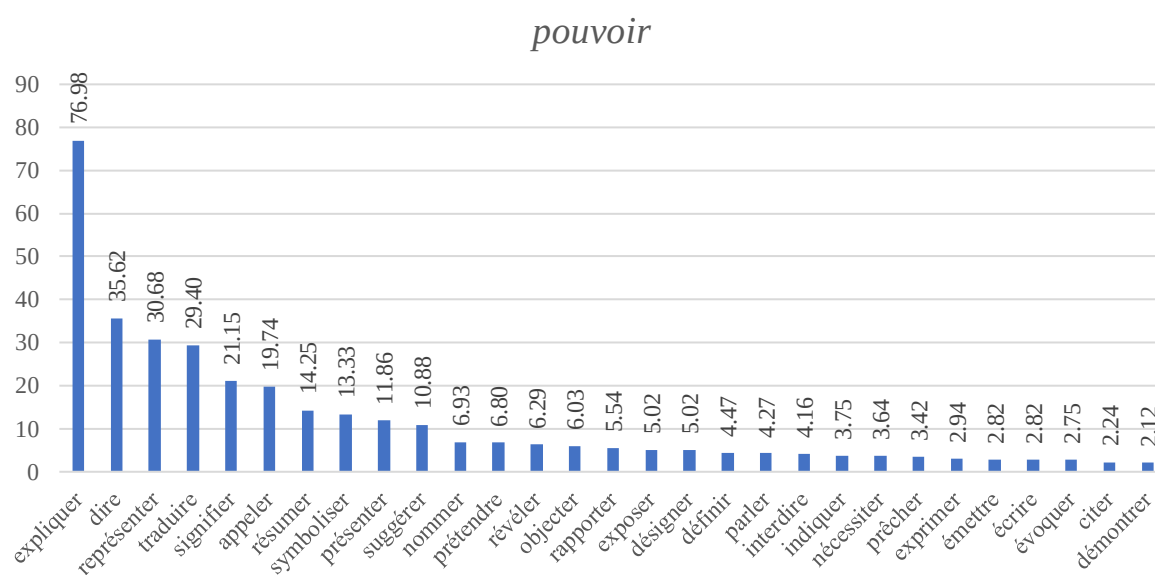


Figure 69 : Les 29 verbes de communication en position de complémentation verbale avec le plus grand indice de spécificité pour *pouvoir* au conditionnel présent dans le corpus WIKI.

Au conditionnel, et avec cette catégorie de verbes, *pouvoir* présente des effets de sens tirés des valeurs modales épistémiques. Il permet, par exemple, de dénoter une *hypothèse* sur une théorie comme nous l'avons déjà soulevé pour les sujets nominaux qui précèdent *pouvoir* au conditionnel (7.1.4.10) :

340. « Taraconta » ***pourrait*** ainsi **désigner** les Estoniens comme les adorateurs de Taara. (WIKI)

341. Si cette ordonnance concerna réellement Cervantès et non un homonyme, elle ***pourrait expliquer*** sa fuite en Italie. (WIKI)

Le *pouvoir* marquant une alternative – comme pour les exemples présentés dans la partie sur les formes lemmatisées (7.2.4.2.) – est également une interprétation possible au *conditionnel* :

342. Par exemple la notation ‘chiens’ désigne l’ensemble de tous les chiens : pour prendre un exemple plus mathématique, on ***pourrait écrire*** parfois ‘pairs’ pour l’ensemble des nombres pairs. (WIKI)

Mais, comme pour les formes au présent de l’indicatif, que ce soit pour les formes lemmatisées de corpus mais aussi pour les divers exemples relevés dans LM, *pouvoir* permet de mettre en avant un fait notable général :

343. On ***pourrait dire*** qu’il s’agit déjà là du principe fondamental de la pensée scientifique, même si les méthodes de la recherche se sont considérablement transformées depuis lors. (WIKI)
344. L’histoire n’est pas racontée dans l’ordre chronologique. Le roman est constitué de sept chapitres que l’on ***pourrait résumer*** comme suit : L’ordre chronologique est 2, 3, 4, 5, 6, 1, 7. (WIKI)
345. On ***pourrait résumer*** les westerns de Leone par la violence du scénario, la musique tonitruante et des acteurs venus de série B américaine. (WIKI)

Ici, ce ne sont pas des énoncés qui reprennent des effets de sens tirés du NS de *pouvoir* : il n’y a pas ici une *possibilité* de *dire* ou de *résumer* puisque la locutrice le fait réellement. *Pouvoir* est ici utilisé comme distanciateur d’un *dire*.

Ainsi, *pouvoir* permet d’exprimer de nombreux usages, relevant des catégories traditionnelles de la modalité, notamment de la valeur modale épistémique, et d’autres où il se désémantise et agit comme un introducteur adoucissant d’un *dire*.

7.2.4.9. Les complémentations verbales de type verbes de serrer/saisir/posséder et don/privation pour devoir au conditionnel présent

Comme nous l’observions également dans LM, *devoir* se combine avec des verbes d’actions plus concrètes que les autres modaux, qui s’associent majoritairement avec les verbes de *communication* et les verbes *psychologiques*. Au conditionnel dans WIKI ce sont donc les verbes de type *serrer/saisir/posséder* et ceux de *don/privation* qui arrivent en tête de liste :

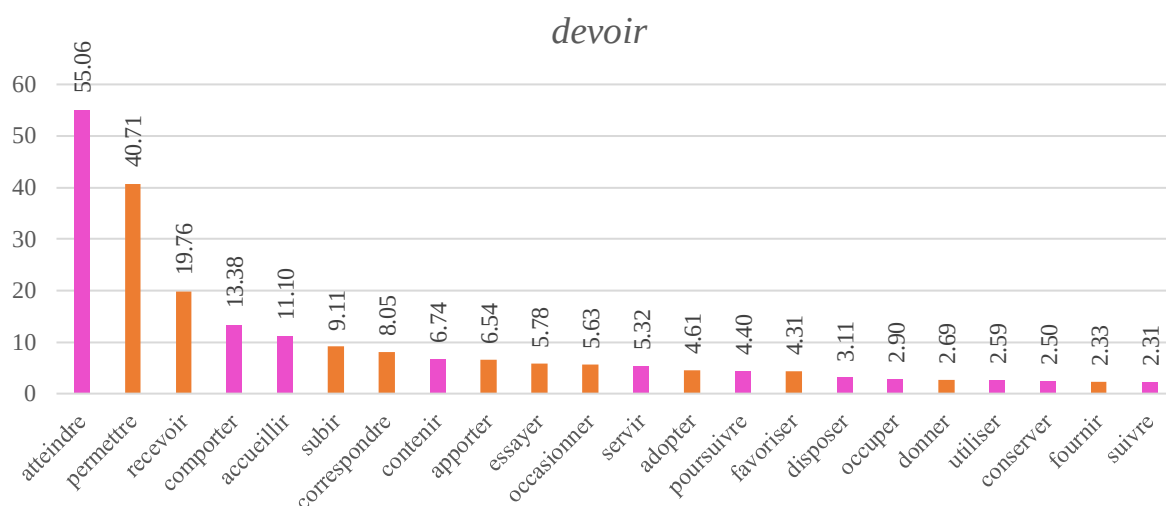


Figure 70 : Les 22 verbes de type saisir/serrer/posséder et de don/privation en position de complémentation verbale avec le plus grand indice de spécificité pour *devoir* au conditionnel présent dans le corpus WIKI.

Avec ces verbes, et par le mode du conditionnel, *devoir* permet à nouveau de présenter des *prédictions* :

346. L’urbanisation de la zone Grimesnil / Monturbet, prévue ces prochaines années, **devrait** logiquement **apporter** un surcroît de population, sans que personne ne sache si celui-ci sera suffisant pour stopper l’hémorragie démographique. (WIKI)
347. Le bâtiment de la gare **devrait subir** également des modifications surtout au niveau du hall pour l’agrandir, l’actuel étant jugé trop étroit. (WIKI)
348. Elle **devrait recevoir** ses premiers Boeing 787 début septembre 2012 et ce jusqu’en 2018. (WIKI)
349. La fusion réalisée en totalité en actions, **devrait permettre** aux actionnaires de American Airlines de détenir 72 % du nouveau groupe. (WIKI)

La *prédiction* dans WIKI peut s’expliquer par sa nature collaborative et en ligne, qui en fait un texte pouvant encore évoluer dans le temps et relater des faits que ne se sont pas encore produits. Cette encyclopédie n’est pas « finie » comme peuvent l’être des encyclopédies traditionnelles. Rappelons que même si ces usages peuvent être considérés comme des expressions de la valeur

modale épistémique, ils peuvent se rattacher au NS de *nécessité*, car avec ce que nous savons, il y a *nécessité* de croire que tel ou tel évènement va survenir¹⁸⁷.

Dans d'autres occurrences, le NS de *nécessité* est également percevable même si dans les exemples suivants nous n'avons pas des cas de *prédiction* mais plutôt d'*obligation* ou de *nécessité* adoucie à l'aide du mode du conditionnel :

- 350. En réalité, ce nom de personne **devrait comporter** un astérisque car il n'est pas attesté. (WIKI)
- 351. Les listes, qu'elles soient énumératives ou non, ne **devraient comporter** aucun élément non terminé par une ponctuation, ne serait-ce qu'une virgule ou un point-virgule. (WIKI)
- 352. [...] les développeurs OpenBSD **devraient essayer** de faire ce qui est correct, propre et sécurisé, même au détriment de la facilité d'utilisation, de la vitesse ou des fonctionnalités. (WIKI)

Nous pouvons ainsi observer que les usages de *devoir* se rattachent à son NS traditionnellement admis, celui de la nécessité. Toutefois, le mode conditionnel « adoucit » cette nécessité dans les énoncés que nous avons examinés. Il est intéressant de noter que la lecture de la valeur modale épistémique, relevée avec les sujets les plus fréquemment employés avec *devoir* au conditionnel, n'est pas spontanée pour ce verbe à ce mode lorsqu'on observe de plus près quelles sont les complémentations verbales les plus communes pour lui.

7.2.4.10. Les complémentations verbales de type verbes de union/réunion pour falloir au conditionnel présent

Concernant *falloir* au conditionnel présent dans WIKI, ce sont les verbes de type *union/réunion* qui sont les plus représentés parmi ses cooccurents spécifiques. Seulement deux verbes de ce type arrivent en cooccurrent spécifique de *falloir*, il s'agit d'*ajouter* (avec un indice de 14,41) et de *distinguer* (avec un indice de 4,12). Dans les occurrences avec ces verbes, *falloir* au conditionnel permet d'ajouter de manière adoucie une information à prendre en compte afin de mieux comprendre le texte (353) mais aussi un conseil pour réussir une certaine tâche (354) :

- 353. Dans l'histoire de la pédagogie, il **faudrait distinguer** méthodes, systèmes, mouvements, démarches, dispositifs, modèles, approches, pratiques. (WIKI)

¹⁸⁷ A ce propos, revoir l'explication de De Saussure (2017) dans la partie 4.3.10 qui explique comment le NS de *nécessité* peut être retrouvé dans les usages épistémiques de ce type pour *devoir*.

354. Toujours selon le CGDD, si l'on voulait aussi dépolluer les nappes, il **faudrait** encore **ajouter** une somme comprise entre 32 et 105 milliards d'euros (dont seulement 7 milliards d'euros pour le respect de la Directive Eaux souterraines). (WIKI)

Dans ce dernier exemple, c'est tant la forme au conditionnel que la nature impersonnelle de *falloir* qui permet à la locutrice de mettre une certaine distance sur le conseil prodigué.

Cependant, ces usages ne sont pas fréquents et ne permettent ainsi pas de tirer des conclusions plus générales sur un quelconque usage particulier de *falloir* au conditionnel avec ce type de verbe.

7.2.4.11. Les complémentations verbales de type verbes de communication pour vouloir au conditionnel présent

Vouloir s'associe, une fois de plus, majoritairement à des verbes de *communication* avec les paramètres étudiés ici. Dans cette configuration, cela représente seulement trois verbes : *dire* (indice de 21,54), *plaider* (indice de 6,96) et *expliquer* (indice de 4,03).

Vouloir au conditionnel suivi de *dire* est grammaticalisé dans presque 90% des cas :

355. Lannilis **voudrait** donc **dire** le « lieu consacré » ou l'« ermitage de l'église », sous-entendu près de ou lié à l'église. (WIKI)
356. Pour Servius « Cerbère » serait un dérivé de « creoberos » qui **voudrait dire** « dévoreur de chair ». (WIKI)

Ici, avec le conditionnel, il y a *supposition* sur la *signification* d'un terme. Quand la locution ne se fige pas, c'est à nouveau parce que le sujet grammatical représente un *animé*, en l'occurrence dans l'exemple suivant le pronom *je* :

357. À tout le personnel d'aérospatiale, des ingénieurs aux techniciens et à tous les travailleurs, je **voudrais dire**, pour la joie qu'ils me donnent aujourd'hui, de tout cœur merci. (WIKI)

Le NS d'*intention* est donc récupérable comme c'est le cas avec *plaider* et *expliquer* :

358. C'est bien là le défaut de la science : elle **voudrait** tout **expliquer** ; et quand il lui est impossible d'expliquer, elle déclare qu'il n'y a rien à expliquer. (WIKI)

359. Mais le développement durable est aussi l'objet de nombreuses critiques. Luc Ferry, par exemple, se demande « qui **voudrait plaider** pour un “développement intenable” ! Évidemment personne ! [...] L'expression chante plus qu'elle ne parle ». (WIKI)

Vouloir confirme donc ici les propriétés qu'on lui a attribué jusqu'à présent avec les verbes de *communication* : il se grammaticalise dans la plupart des cas avec *dire* et avec les autres verbes de cette catégorie, il exprime son NS d'*intention/volonté* de manière plus adoucie au conditionnel.

7.2.5. L'AFC et l'analyse arborée appliquée aux complémentations verbales sous forme de verbe à l'infinitif

Comme pour les sujets nominaux, nous allons ici étudier les propriétés combinatoires des verbes modaux avec leurs complémentations verbales sous forme de verbes à l'infinitif à l'aide des AFC et des analyses arborées. Cette nouvelle couche d'analyse nous permettra de mettre en parallèle les résultats relevés dans les parties précédentes et de comparer ainsi les verbes modaux de manière plus directe. Au vu des remarques que nous avons soulevées lors de l'analyse de leurs profils combinatoires et des exemples choisis sur la base de leur indice de spécificité, il sera intéressant de voir si les rapprochements et éloignements repérés sont également visibles dans les AFC et l'analyse arborée, comme par exemple le parallèle entre *pouvoir* et *vouloir* et leurs associations importantes avec les verbes de *dire* ou encore le fait que *devoir* s'éloigne des autres modaux en étant plus souvent usité avec des verbes d'actions concrètes.

7.2.5.1. Les résultats dans LM pour les formes lemmatisées

La première AFC que nous allons analyser est celle des modaux sous leur forme lemmatisée dans LM :

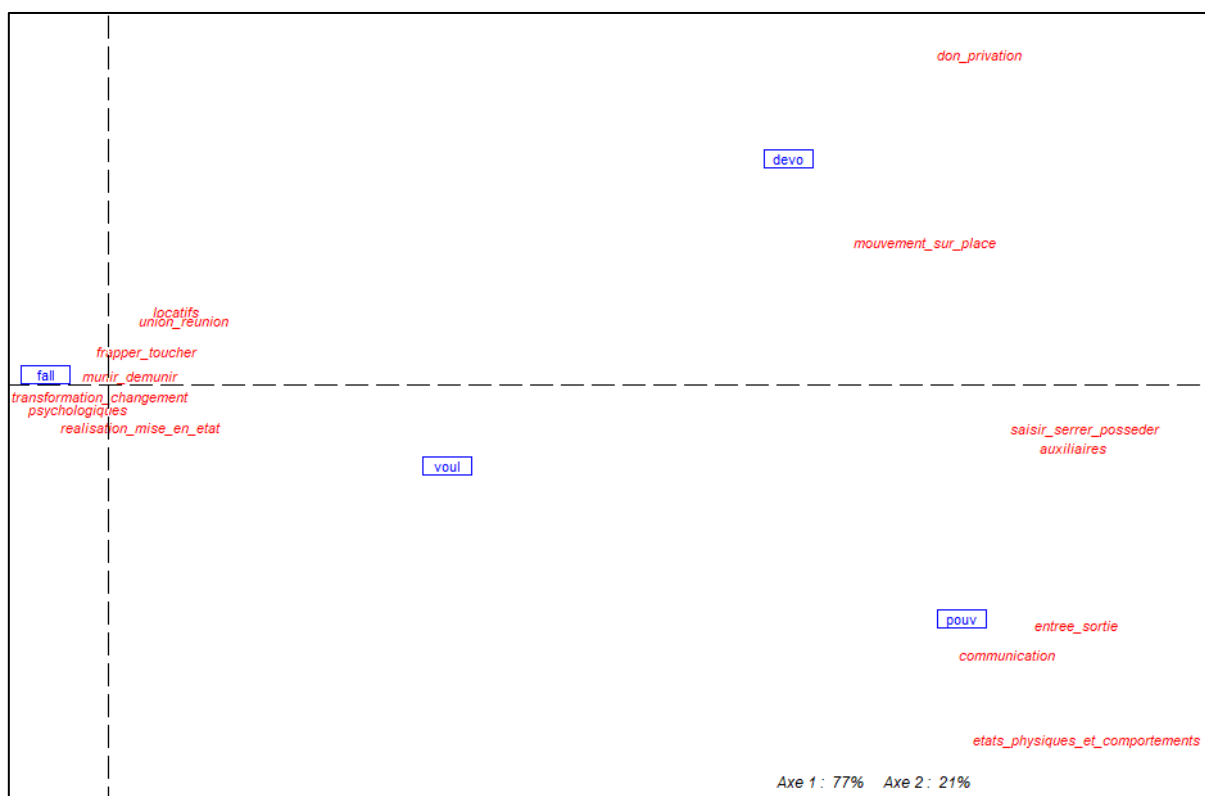


Figure 71 : AFC des verbes sous leurs formes lemmatisées avec les catégories sémantiques de Dubois et Dubois-Charlier (1997) des verbes en position complémentation verbale dans LM.

L'opposition principale sur l'axe 1 concerne *pouvoir* et *falloir*, ce qui paraît étonnant au vu de leurs usages similaires que nous avons pu relever dans les différents exemples. Notons toutefois que *falloir* se trouve à proximité des verbes *psychologiques* et *pouvoir* des verbes de *communication*. Il semble ainsi que même si ces deux modaux présentent des usages similaires avec ces deux catégories, l'un « préfère » les verbes de *pensée* et l'autre les verbes de *parole*. Sur ce même axe horizontal – qui contient 77% des informations représentées – *pouvoir* et *devoir* sont proches. Cela confirme une certaine similarité dans leurs profils combinatoires pour les formes lemmatisées dans ce corpus (7.2.3.1.). Notons toutefois que *pouvoir*, tout comme *falloir*, se retrouve avec beaucoup de catégories verbales autour de lui. Ces catégories sont très diversifiées avec des actions plus « passives » (verbes de *communication* ou *auxiliaires*) et d'autres plus « actives » (*entrée/sortie* et *saisir/serrer/posséder*), ce qui démontre ici sa nature hautement polysémique qui l'amène à s'associer avec beaucoup de verbes différents. *Devoir*, quant à lui, est entouré de catégories sémantiques qui dénotent des actions plus concrètes, comme les verbes de *don/privation* ou encore ceux de *mouvement sur place*.

Concernant l'axe 2, la distinction se fait surtout entre *pouvoir* et *devoir*. Ainsi, même si des éléments les rassemblent sur l'axe 1, il semble qu'il y a toutefois des différences notables entre ces deux verbes. En effet, contrairement aux sujets, nous avons noté moins de ressemblances entre ces deux verbes concernant le type de verbe qu'ils modalisent. L'axe vertical comprend

bien moins d'information mais il est intéressant de relever que sur celui-ci, *devoir* et *falloir* sont relativement proches. *Falloir* se trouve également à proximité de *vouloir*, comme c'était le cas sur l'axe 1, ce qui confirme leur proximité relativement importante. *Vouloir*, quant à lui, avoisine *pouvoir* ce qui renforce le lien soulevé déjà par leur position dans le même cadran.

En mettant en parallèle les informations que nous avons pu relever pour les deux axes, nous pouvons conclure que *pouvoir* et *devoir* partagent des propriétés communes, et cela peut être dû à leurs usages épistémiques (ceci expliquerait ainsi leurs positions sur l'axe 1). Cependant, ils sont éloignés sur l'axe 2, qui pourrait davantage représenter leurs NS, ce qui justifierait ainsi également la proximité entre *devoir* et *falloir* sur cet axe.

Concernant le centre de l'AFC, nous y retrouvons à nouveau *falloir*, ce qui était également le cas pour les AFC représentant leurs associations avec les sujets. Une fois de plus, il ne semble ainsi pas se démarquer des autres verbes par rapport à ses propriétés combinatoires. Cela était moins étonnant pour les sujets, étant donné que ceux-ci étaient bien moins présents pour ce modal que pour les autres. Nous pouvons ici émettre l'hypothèse que c'est sa nature de verbe impersonnel qui le rend moins sensible à une attraction avec des verbes en particulier mais c'est bien le seul point sur lequel *falloir* se différencie des autres verbes ici.

Le deuxième modal le plus proche du centre est *vouloir*. Cependant, contrairement à tous les autres modaux sur ce graphique, apparaît un certain « vide » autour de lui. Il ne semble ainsi pas attirer une catégorie de verbe en particulier en position de complémentation verbale. Il paraît plutôt « tiraillé » entre les trois autres modaux que nous analysons ici : ainsi, il est relativement proche du centre et donc de *falloir* – ce qui le rend relativement « banal » par rapport à ses propriétés combinatoires avec les verbes en position de complémentation verbale – mais il se situe également dans le même cadran que *pouvoir* – peut-être pour ses associations avec les verbes de *communication* qui se trouvent dans ce même cadran – tout en étant à proximité de *devoir* sur l'axe horizontal – avec qui il partage donc également certaines propriétés combinatoires.

Ces proximités et éloignements entre les modaux se confirment dans l'analyse arborée de ces mêmes chiffres :

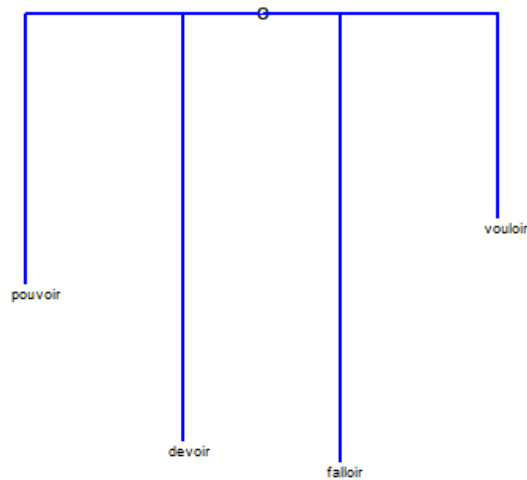


Figure 72 : Analyse arborée des modaux sous leurs formes lemmatisées avec les catégories sémantiques de Dubois et Dubois-Charlier (1997) des verbes en position de complémentation verbale dans LM.

Le nœud principal sépare ainsi *pouvoir* et *devoir* d'un côté et *falloir* et *vouloir* de l'autre, ce qui représente les proximités et éloignements dénotés sur l'axe 1 de l'AFC. Rappelons que les séparations par nœuds sont plus importantes en termes d'informations contenues que la longueur des branches. Nous pouvons ainsi admettre que les points communs entre *pouvoir* et *devoir* sont donc plus importants que ceux qui les séparent.

Comme pour les sujets, nous allons également observer si les rapprochements et éloignements entre ces verbes peuvent être affinés à l'aide d'une mise en perspective par temps et mode verbal.

7.2.5.2. Les résultats dans LM pour les formes au présent de l'indicatif et au conditionnel présent

Contrairement aux AFC par temps et mode verbal analysés pour les sujets nominaux, nous avons ici pris en compte les quatre modaux pour en étudier les similarités et disparités, puisque *falloir* compte ici suffisamment d'occurrences avec les complémentations verbales pour être considéré :

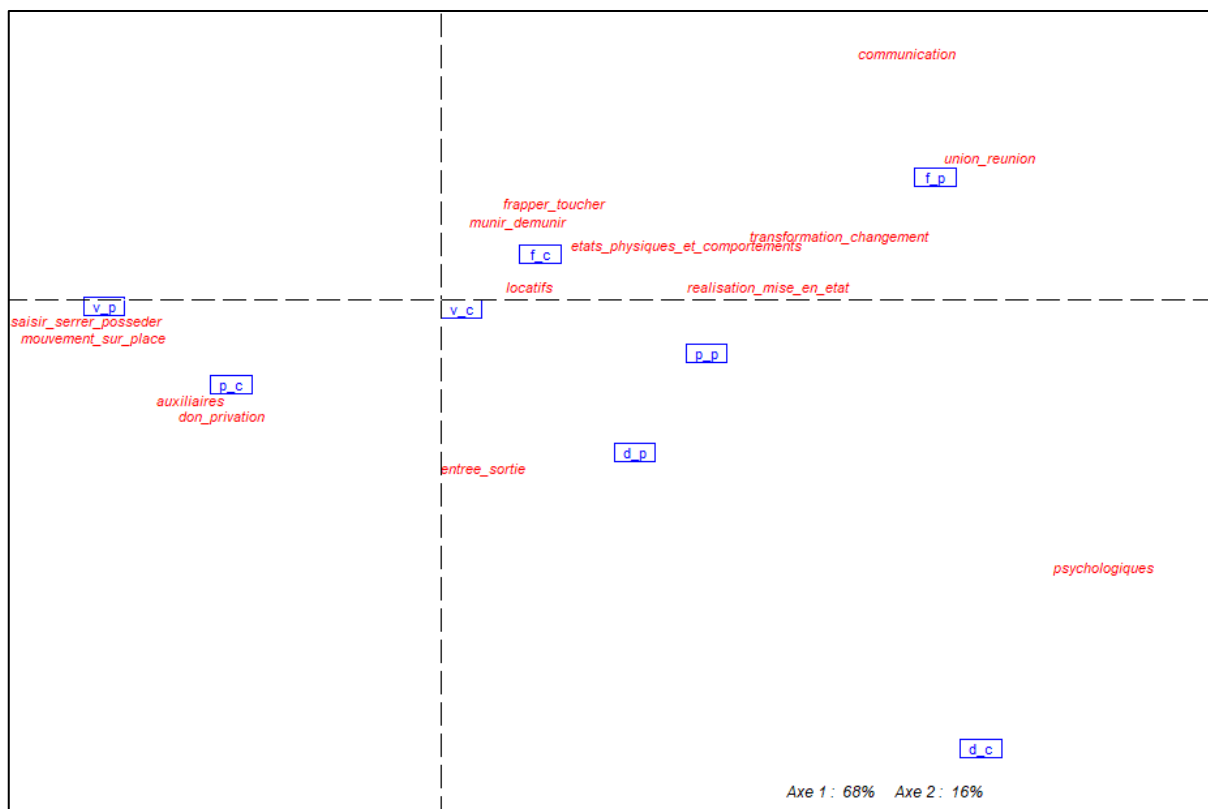


Figure 73 : AFC des modaux au présent de l'indicatif et au conditionnel présent avec les catégories sémantiques de Dubois et Dubois-Charlier (1997) des verbes en position de complémentation verbale dans LM.

Sur l'axe 1, se trouvent d'un côté *vouloir* au présent et *pouvoir* au conditionnel et de l'autre *falloir* au présent et *devoir* au conditionnel. Cette opposition démontre qu'il y a une certaine proximité à noter dans les usages entre ces deux regroupements et ce sont ces ressemblances qui les opposent. *Devoir* au conditionnel se trouve dans le même cadran en bas à droite que *devoir* au présent, ce qui, comme pour *falloir*, dénote une certaine stabilité de son NS vu que, peu importe le temps et le mode verbal auquel il est conjugué, les propriétés combinatoires sont similaires. *Pouvoir* au présent se trouve également à proximité de *devoir* au présent, ce qui démontre une fois de plus une similarité entre ces deux verbes malgré leurs NS traditionnellement opposés dans les études sur la modalité.

Sur l'axe 2, cependant, *devoir* au conditionnel se distingue de toutes les autres formes et en particulier de *falloir* au présent dont il était proche sur l'axe 1. Toutefois, ce qui est intéressant de noter sur l'axe 2 est que les deux formes de chaque modal sont à chaque fois les plus proches de ce point de vue. Il semble donc que leurs NS semblent toujours rassembler d'une certaine manière les différents usages d'un même modal. Ceci est renforcé pour *falloir* si nous adoptons la lecture par cadran. Dans le cadran en haut à droite sont localisées les deux formes de *falloir*, ce qui démontre une certaine régularité dans ses usages, peu importe le temps ou le mode verbal auquel il est conjugué. Il est d'ailleurs entouré d'énormément de catégories verbales, ce qui

était le cas également pour sa forme lemmatisée lorsqu'il était au centre des deux axes. Il ne semble donc pas y avoir une de ces catégories qui singularise plus qu'une autre son comportement.

Proche du centre, nous retrouvons *vouloir* au conditionnel avec, à proximité, *falloir* au conditionnel également. Dans toutes les AFC que nous avons analysées jusqu'à présent, ces deux verbes se retrouvent systématiquement au centre (ou proches du centre) des deux axes, ce qui démontre qu'ils ne se démarquent pas autant que *pouvoir* et *devoir* sur la base de leurs propriétés combinatoires. Ils ont ainsi des cotextes d'apparition moins différenciés que les deux autres modaux dont le NS est peut-être plus marqué et permet ainsi des usages plus spécifiques. Cette proximité avec le centre des deux axes peut également être due à leurs propriétés syntaxiques : en effet, ces deux modaux n'ont pas nécessairement besoin d'un verbe à l'infinitif afin de dénoter un effet de sens (4.5.1.).

L'analyse arborée vient toutefois quelque peu nuancer ces propos :

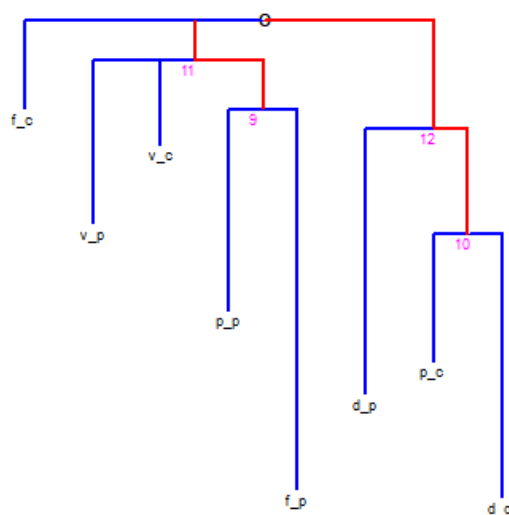


Figure 74 : Analyse arborée des modaux au présent de l'indicatif et au conditionnel présent avec les catégories sémantiques des verbes de Dubois et Dubois-Charlier (1997) en position de complémentation verbale dans LM.

Nous retrouvons, du côté droit du nœud, les deux formes de *devoir* avec *pouvoir* au conditionnel. *Pouvoir* et *devoir* au conditionnel se retrouvent tous deux regroupés dans un embranchement commun, ce qui rejoint les réflexions générales que nous avons déjà émises à leur propos : au conditionnel, ces deux verbes sont proches et ceci pourrait être dû au fait qu'à ce mode verbal ils dénotent principalement des valeurs modales épistémiques. Ceci pourrait aussi expliquer pourquoi *devoir* au présent de l'indicatif se retrouve également de ce côté-là du nœud puisqu'il peut véhiculer l'effet de sens de prédiction dans le corpus de presse. Ce regroupement des deux formes de *devoir* d'un même côté du nœud dénote, encore une fois, d'une certaine stabilité de son NS.

De l'autre côté du nœud, nous retrouvons toutes les autres formes étudiées. Tout à gauche, et isolé des autres items, se situe *falloir* au conditionnel. C'est également l'item qui se retrouve le plus proche du nœud central. Cette proximité avec le nœud peut donc s'expliquer avec une certaine « banalité » dans son comportement par rapport aux autres formes, ce qui avait déjà été relevé dans l'AFC dans la figure 75.

A gauche de ce nœud principal, nous retrouvons deux autres regroupements : les deux formes de *vouloir* dans un même embranchement puis *pouvoir* et *falloir* au présent dans un autre. *Vouloir* semble ainsi avoir un comportement stable quel que soit le temps ou le mode verbal auquel il est conjugué, ce qui était déjà perceptible avec les profils combinatoires que nous avons relevés. Quant à *pouvoir* et *falloir* au présent de l'indicatif, ils sont rassemblés ici dans un même embranchement ce qui signifie que leurs similarités que nous avons relevées dans la section 7.2.3.5. par rapport à leurs associations avec les verbes *psychologiques* est un point commun assez significatif pour les rassembler ici. Les ressemblances entre ces deux formes sont donc plus importantes que celles que possèdent les deux formes de *falloir* ou celles de *pouvoir*. D'ailleurs, les deux formes de *pouvoir* se retrouvent séparées par le nœud principal : nous avons donc d'un côté ses usages plutôt de l'ordre de la valeur modale épistémique – avec *pouvoir* au conditionnel présent – et de l'autre les formes plus grammaticalisées, lorsqu'il agit comme un marqueur qui attire l'attention sur un fait général – *pouvoir* au présent de l'indicatif avec les verbes *psychologiques*. Rappelons que *pouvoir* et *falloir* n'ont été, à ce jour, jamais regroupés dans une étude ou un classement des usages de ces verbes modaux, ce qui démontre que les méthodes que nous employons ici permettent de soulever des propriétés importantes de ces verbes encore méconnues.

7.2.5.3. Les résultats dans WIKI pour les formes lemmatisées

En ce qui concerne les résultats dans WIKI avec les complémentations verbales sous forme de verbes à l'infinitif, les rapprochements et éloignements entre les quatre verbes modaux que nous étudions ici sont relativement différents de ceux relevés dans LM :

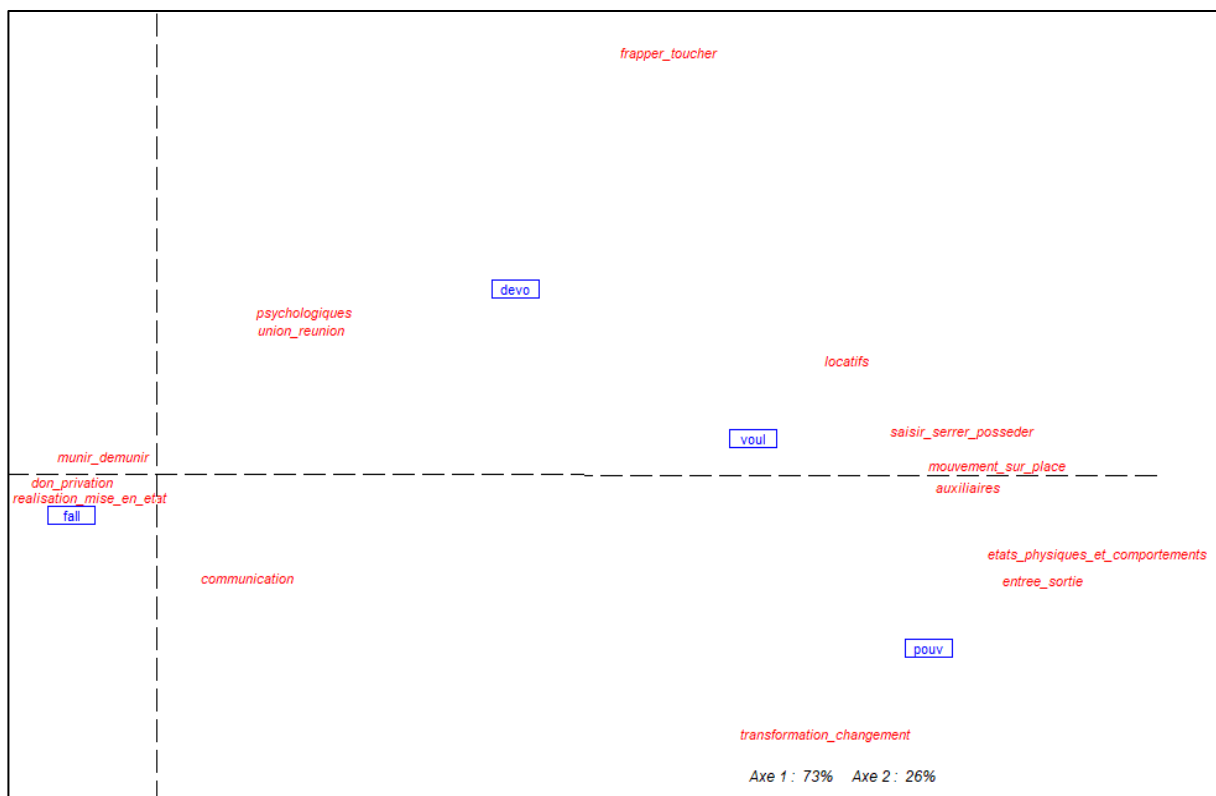


Figure 75 : AFC des verbes sous leurs formes lemmatisées avec les catégories sémantiques de Dubois et Dubois-Charlier (1997) des verbes en position complémentation verbale dans WIKI.

Si on lit le graphique de gauche à droite en nous focalisant ainsi sur l'axe 1 – qui comporte 73% des informations – nous pouvons nous apercevoir que *falloir* et *pouvoir* s'opposent drastiquement alors que nous avons dénoté des usages similaires comme marqueur de focalisation sur des faits généraux. Nous nous apercevons également que *devoir* est le deuxième verbe le plus proche du centre et donc le modal le plus similaire à *falloir*. Leurs NS de *contrainte* et/ou de *nécessité* qu'on leur attribue généralement semblent ainsi ici plus marqués que dans le corpus de presse où leurs usages étaient plus différenciés.

Sur l'axe 2, ce sont à nouveau *pouvoir* et *devoir* qui s'opposent, ce qui démontre que leurs différences sont relativement importantes puisque cette distance apparaît dans plusieurs de nos AFC. Dans la presse, cette différence était moins marquée. Ceci peut être dû au fait que la valeur modale épistémique véhiculée par ces deux verbes est moins présente dans le discours encyclopédique que dans le corpus de presse. Sur cet axe, *devoir* et *vouloir* se retrouvent également dans le même cadran. Ce rassemblement peut être dû aux catégories de verbes plutôt « actifs » auxquelles ils se trouvent associés : *frapper/toucher*, *locatifs*, *saisir/serrer/posséder*, etc. Ce type de verbe pourrait se trouver plus fréquemment associé à ces modaux car ils attirent majoritairement des sujets *animés* dans ce corpus (cf. 7.1.4.1. et 7.1.4.2.). Ces sujets se trouvent plus facilement associés à ce type de verbes. Ainsi, c'est donc leurs NS de *nécessité* et

d'*intention/volonté* qui les regroupent ici. *Vouloir* se trouve cependant aussi à proximité de *pouvoir* sur l'axe 1. Ces deux modaux démontraient déjà d'une certaine proximité dans l'AFC des formes lemmatisées dans LM (cf. figure 73 dans la section 7.2.5.1.) car ils se retrouvaient dans le même cadran. Cette proximité semble ici se confirmer dans un autre type de texte. Dans leurs profils combinatoires, nous avons observé qu'ils avaient tous deux tendance à attirer majoritairement des verbes de *communication* mais il semble qu'ils aient d'autres points communs importants puisqu'ils se retrouvent à proximité des verbes de *mouvement sur place*, des *auxiliaires* ou de verbes d'*états physiques/comportements*. Si nous observons toutefois leurs positions respectives par rapport à l'axe 2, les verbes de *communication* se trouvent bel et bien entre ces deux modaux.

Comme dans les autres AFC, nous retrouvons à nouveau *falloir* au centre des deux axes, ce qui confirme une propriété stable de ce modal : il ne se différencie pas des autres modaux par rapport à ses associations. Cependant, dans WIKI, contrairement à ce que nous avons pu observer pour le corpus de presse, il y a moins de catégories sémantiques verbales qui le rejoignent au centre.

L'analyse arborée vient quelque peu nuancer ces propos :

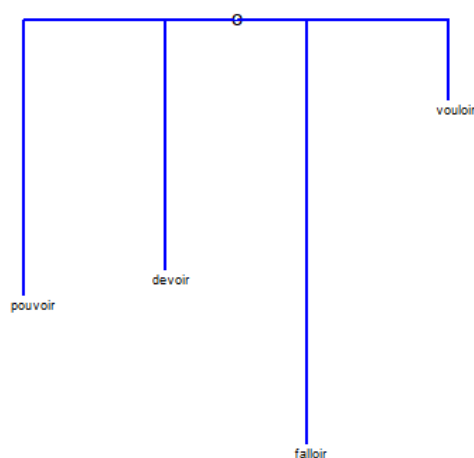


Figure 76 : Analyse arborée des verbes sous leurs formes lemmatisées avec les catégories sémantiques des verbes de Dubois et Dubois-Charlier (1997) en position de complémentation verbale dans WIKI.

Comme pour l'analyse arborée des formes lemmatisées dans LM (cf. figure 74 dans la section 7.2.5.1.), *pouvoir* et *devoir* se retrouvent à gauche du nœud et *falloir* et *vouloir* rassemblés de l'autre côté. Les points communs partagés entre *pouvoir* et *devoir* semblent donc, une fois de plus, plus importants que leurs divergences. Il est également possible que la plus grande fréquence de ces deux modaux leur procure davantage de possibilités de rencontres avec d'autres verbes et donc plus de points communs, ce qui peut ainsi les rassembler systématiquement dans les analyses arborées fondées sur les lemmes. C'est pourquoi une

analyse avec les formes du présent de l'indicatif et du conditionnel présent permet d'affiner ces résultats.

7.2.5.4. Les résultats dans WIKI pour les formes au présent de l'indicatif et au conditionnel présent

Les verbes modaux semblent avoir un usage différencié selon le type de texte dans lequel ils sont usités, du moins, leurs ressemblances et divergences ne sont pas tout à fait les mêmes entre notre corpus de presse et le corpus encyclopédique :

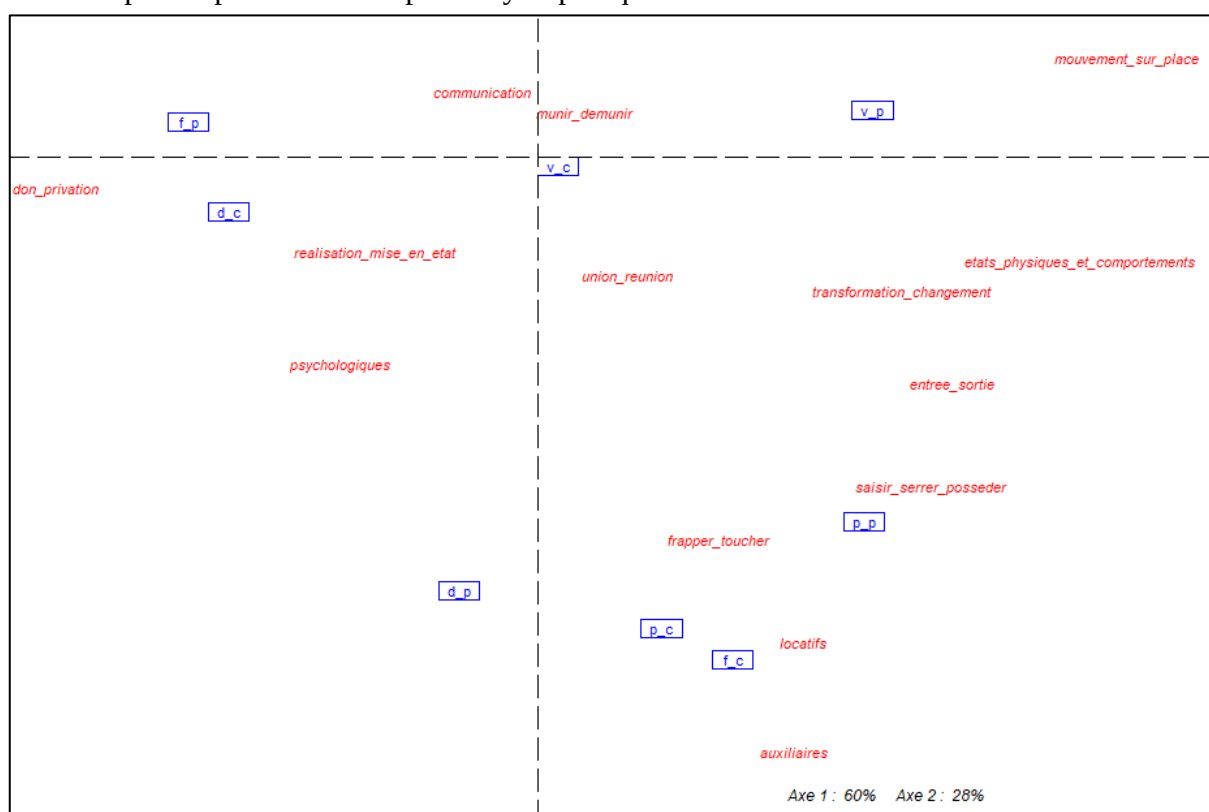


Figure 77 : AFC des modaux au présent de l'indicatif et au conditionnel présent avec les catégories sémantiques des verbes de Dubois et Dubois-Charlier (1997) en position de complémentation verbale dans WIKI.

Sur l'axe 1, nous retrouvons, comme pour le corpus LM, *falloir* au présent de l'indicatif d'un côté, accompagné de *devoir* au conditionnel présent, et de l'autre, *vouloir* au présent de l'indicatif, mais cette fois-ci accompagné de *pouvoir* à ce même temps et mode. La proximité entre *falloir* et *devoir* au présent de l'indicatif semble ainsi un facteur stable dans nos deux corpus. *Devoir* au conditionnel permet de prendre une certaine distance sur la *nécessité*, prise de distance que *falloir*, qui par sa nature de verbe impersonnel, instaure seul. Cette proximité est également confirmée sur l'axe 2.

Sur l'axe 2, *falloir* au conditionnel présent s'oppose à *vouloir* au présent de l'indicatif, voire à *falloir* sous cette même forme. Contrairement à LM, *falloir* comporte donc des usages distincts selon le temps et le mode auquel il est conjugué.

Au centre, nous retrouvons, comme dans le corpus LM, la forme au conditionnel de *vouloir*. En haut à droite, c'est également le verbe *vouloir* qui apparaît mais cette fois-ci au présent de l'indicatif. Il se retrouve seul dans ce cadran. Cependant, en observant de plus près sa position sur l'axe 1, nous remarquons qu'il se situe sur le même point que *pouvoir* au présent. Cette proximité pourrait ainsi en partie expliquer pourquoi ils étaient déjà environnants sous leur forme lemmatisée.

Pouvoir au présent se localise dans le cadran en bas à droite avec sa forme au conditionnel, ce qui démontre dans WIKI une plus grande régularité dans ses usages selon le temps et le mode verbal que dans LM où ces deux formes étaient relativement éloignées. Ces deux formes de *pouvoir* se retrouvent également avec *falloir* au conditionnel dans ce cadran. Ce sont les formes au conditionnel qui sont le plus proches, ce qui pourrait donc être dû au mode qui joue ici un grand rôle dans les cotextes dans lesquels ils apparaissent. Dans ce cadran figurent aussi beaucoup de catégories sémantiques de verbes des complémentations verbales. Ceci appuie une fois de plus la capacité à *pouvoir* de s'associer avec beaucoup de types de verbes différents dû à sa polysémie de plus grande envergure que pour les autres modaux.

Dans le cadran en bas à gauche, on retrouve cette fois-ci les deux formes de *devoir* avec, entre autres, la catégorie des verbes de *don/privation* avec laquelle il semble avoir une affinité particulière, quel que soit le temps ou le mode verbal auquel il est usité. Cette proximité entre ces deux formes atteste également de la stabilité du NS de *devoir* qui ne semble pas avoir des effets de sens ou des valeurs modales trop distinctes selon les formes qu'il peut prendre.

Les verbes les plus réguliers dans leurs NS semblent donc être *pouvoir* et *vouloir* dont les formes au présent de l'indicatif et au conditionnel présent sont proches tant sur l'axe 1 que 2, alors que pour *devoir* et *falloir*, les usages semblent un peu plus différenciés selon le temps et le mode verbal auquel ils sont utilisés.

Cependant, il semble bien que les deux formes de chaque verbe soient finalement plus proches que ne le laissent paraître l'AFC si l'on observe l'analyse arborée de ces mêmes données ci-dessous :

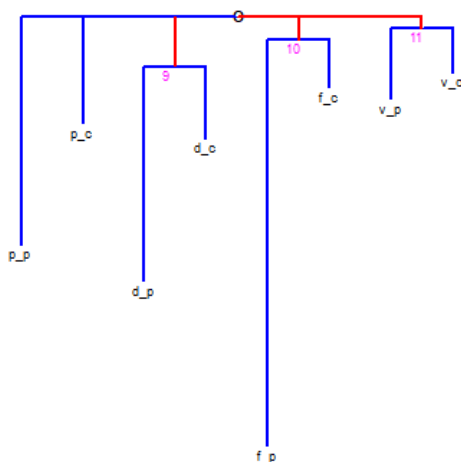


Figure 78 : Analyse arborée des modaux au présent de l'indicatif et au conditionnel présent avec les catégories sémantiques des verbes de Dubois et Dubois-Charlier (1997) en position de complémentation verbale dans WIKI.

Le nœud principal sépare à nouveau *pouvoir* et *devoir* d'un côté, et *falloir* et *vouloir* de l'autre. Les deux formes de chaque modal sont également regroupées dans un même nœud à chaque fois, ce qui diffère fortement de ce que nous avons obtenu comme graphique dans le corpus LM.

En somme, il semblerait que, dans WIKI, les NS des modaux soient plus distincts entre eux. Ceci se remarque par le fait que, peu importe le temps ou le mode verbal auquel ces verbes sont conjugués, il y a plus de similarités entre deux formes d'un même verbe qu'entre des formes de verbes différents, ce qui n'était pas le cas dans LM.

Synthèse sur les associations entre les verbes modaux et leur complémentation verbale

Comme pour les noms qui se trouvaient en position sujet, nous avons choisi de catégoriser les verbes des complémentations verbales dans des groupements sémantiques que nous avons empruntés à Dubois et Dubois-Charlier (1997). Les regroupements sémantiques pour les verbes sont un peu plus complexes à délimiter que pour les noms car les verbes sont davantage polysémiques et n'ont pas une correspondance *signifiant-signifié* aussi évidente que les noms. Cette classification nous a toutefois permis de mieux analyser nos données, de la même manière que nous l'avons fait avec les noms.

Nous avons rappelé en introduction de cette partie que certaines autrices admettent que les modaux « choisissent » le type de verbe qui suit. Ainsi, leurs comportements pourraient être distingués sur cette complémentation verbale.

Même si l'importance de cette complémentation verbale a été soulignée pour déterminer les propriétés sémantiques des verbes modaux, celle-ci n'a pas fait l'objet de beaucoup de réflexions ou de recherches. Les caractéristiques relevées à ce propos étaient d'ailleurs très vagues : comme pour les sujets, on attribue telle ou telle valeur modale ou effet de sens à *pouvoir*, *devoir*, *falloir* ou *vouloir* en fonction d'un type de verbe qui suit, sans détailler grand-chose à ce propos. Les méthodes quantitatives que nous employons dans cette thèse permettent justement de dépasser ces limitations. En effet, les études antérieures se focalisaient souvent sur des catégories de verbes qui ne sont pas les plus fréquemment associées aux modaux, comme nos résultats l'ont montré. En se concentrant sur les associations les plus importantes, nous sommes en mesure de relever des caractéristiques saillantes des propriétés sémantiques des verbes modaux notamment les phénomènes de grammaticalisation qui se sont avérés fréquents pour trois des quatre verbes étudiés ici.

Concernant les résultats en tant que tels, nous avons pu confirmer les phénomènes de grammaticalisation mentionnés dans la littérature mais nous avons aussi pu soulever des propriétés encore peu reconnues pour ces verbes.

Pouvoir se trouve systématiquement associés de manière importante aux verbes de *communication* et aux verbes *psychologiques*. Nous avons pu relever que *pouvoir* permet d'exprimer des utilisations relevant des catégories modales traditionnelles en exprimant des *possibilités* dans le cadre des valeurs modales épistémiques mais aussi des emplois grammaticalisés. Ainsi, avec des verbes de *communication* et *psychologiques*, il peut occuper la fonction de marqueur de « focalisation » puisqu'il attire l'attention de la lectrice ou de l'interlocutrice sur un fait notable, la plupart du temps généralisable, et ce, quel que soit le temps ou le mode verbal auquel il est conjugué. Dans d'autres usages grammaticalisés, il permet aussi d'être marqueur d'une alternative. Concernant les usages plus communs, il s'avère que le mode du conditionnel lui confère plus souvent une valeur modale épistémique ce qui démontre, une fois de plus, l'importance de ce mode verbal pour l'expression de cette valeur modale. Dans ces cas, il permet, par exemple, d'émettre une *hypothèse* qui est donc bien l'expression d'une *possibilité* sur un fait.

Devoir semble être le verbe qui se singularise le plus par rapport à ses associations avec les verbes en position de complémentation verbale. En effet, même si les verbes de *communication* ou encore les verbes *psychologiques* ont aussi leur importance dans les catégories sémantiques

représentées pour ses associations spécifiques, ce modal se combine de manière significative avec des verbes dénotant des actions plus « concrètes ». Cette propriété a été observée dans ses différents profils combinatoires mais aussi confirmés dans les catégories qui gravitent autour de lui dans les AFC. D'ailleurs, même dans l'AFC qui ne prenait en compte que les verbes de *communication* en commun pour les quatre modaux étudiés dans cette thèse (cf. figure 54, section 7.2.3.2.), *devoir* se rapprochait des verbes plus « actifs » de cette catégorie. Quant aux effets de sens et valeurs modales véhiculées, elles sont toutes rattachables au NS de *nécessité* de ce modal. En effet, très souvent il exprime une *nécessité* ou une *obligation* et même lorsqu'on peut lui attribuer un effet de sens de *prédiction*, celle-ci peut se rattacher au NS de *nécessité*. Contrairement à l'expression d'une incertitude avec *pouvoir*, celles véhiculées avec *devoir* sont des prédictions quasi-certaines : il y a une *nécessité* que le fait que l'on énonce survienne avec les connaissances dont nous disposons. *Devoir* est ici le modal qui semble le plus se rattacher à son NS traditionnel car les autres modaux – comme nous l'avons déjà rappelé avec *pouvoir* ci-dessus – qui présentent aussi souvent des emplois grammaticalisés.

D'ailleurs, *falloir* est sans doute le verbe le plus « volatile » de notre sélection. En effet, même si parfois il dénote la valeur modale déontique – qui est le plus habituellement attribué à ce verbe au vu de son NS – il est très souvent grammaticalisé avec les verbes psychologiques : la locution *faut croire* est grammaticalisée dans la grande majorité des cas, et ses usages avec des verbes comme *imaginer* ou encore *savoir* lui confèrent plutôt la valeur d'un marqueur de focalisation. Lorsque ce n'est pas le cas, au conditionnel notamment, il semble qu'il exprime une valeur modale épistémique car il permet de mettre en avant une pensée alternative. En outre, ses comportements sont très stables d'un corpus à un autre, ce qui montre une régularité d'usage de ce verbe. Cependant, au vu du nombre d'emplois grammaticalisés qu'il revête, nous pouvons nous demander si le NS *nécessité/contrainte* est vraiment le NS qui lui correspond le mieux. Nous pouvons même émettre l'hypothèse qu'il possède tellement d'usages grammaticalisés qu'il fonctionne ainsi davantage comme un marqueur de focalisation avec ces verbes *psychologiques* plutôt que comme un modalisateur dénotant une *nécessité*.

Concernant *vouloir*, celui-ci semble partager différentes caractéristiques avec les autres verbes modaux présentés ici. En effet, il peut agir comme un modal véhiculant son NS d'*intention/volonté* – ce qu'il fait d'ailleurs dans une grande majorité des cas – mais le verbe avec lequel il possède l'indice de spécificité le plus important dans la catégorie des verbes de *communication* – catégorie la plus importante pour *vouloir* dans cinq des six profils combinatoires présentés – est *dire* avec lequel il est grammaticalisé dans la grande majorité des cas. Cette grammaticalisation se retrouve dans tous les corpus et dans toutes les formes étudiées

ici, c'est-à-dire, en tant que lemme mais aussi au présent de l'indicatif ainsi qu'au conditionnel présent. Ceci montre une régularité dans son comportement et donc une propriété stable de verbe. Les cas dans lesquels *vouloir dire* n'est pas grammaticalisé sont les occurrences dans lesquelles le sujet est un *animé*. Ceci montre l'importance de prendre en compte la totalité du co(n)texte du verbe pour pouvoir en interpréter l'effet de sens ou la valeur modale.

Quant aux ressemblances et divergences entre ces verbes, nous avons soulevé plusieurs caractéristiques qui n'ont pas encore été documentées dans la littérature à leur sujet. Ainsi, nous avons pu noter que les verbes modaux étaient majoritairement attirés par des verbes de *communication* et verbes *psychologiques*. Au conditionnel, leurs profils combinatoires sont plus différenciés, mais nous avons pu relever dans les exemples observés que leurs effets de sens étaient similaires, se rattachant souvent à des valeurs modales épistémiques.

Les ressemblances que nous avons pu noter et qui n'ont à ce jour pas été encore relevées dans la littérature concernent notamment *pouvoir* et *falloir*. En effet, ces deux modaux partagent une attraction importante avec les verbes *psychologiques*. L'observation plus détaillée de leurs associations avec ces verbes nous a permis de soulever une capacité similaire de ces deux modaux : ils permettent de dénoter des faits généraux et de porter l'attention de la lectrice ou de l'interlocutrice sur ces généralités. Cette propriété peut être associée à *falloir* puisqu'en tant que verbe impersonnel il n'adresse pas sa modalité exprimée à un groupe, à une personne ou un phénomène en particulier. *Pouvoir*, quant à lui, permet d'exprimer ces généralités lorsqu'il est précédé par le pronom *on* avec lequel il semble fréquemment associé lorsqu'il est suivi par les verbes *psychologiques*. Cependant, le sens de cette généralité diffère quelque peu entre ces deux modaux : *falloir* véhicule une généralité qu'il est *nécessaire* de connaître (dû à son NS de *nécessité* ou *contrainte*), alors que *pouvoir* adoucit le besoin d'être au courant du fait énoncé (dû à son NS de *possibilité*).

Pouvoir partage aussi des points communs avec *vouloir*. Dans LM, au conditionnel, ils s'associent tous deux particulièrement avec les verbes de *communication* et verbes *psychologiques*. Cette proximité ressort également de manière significative dans les AFC.

Vouloir partage aussi des caractéristiques communes avec *falloir*. En effet, dans les AFC que nous avons présentées, les deux verbes sont relativement proches entre eux et du centre des deux axes. Cette proximité s'explique par le fait qu'ils se grammaticalisent tous deux aisément avec le verbe *dire*. Un autre trait qu'ils ont en commun est le fait qu'ils sont tous les deux moins fréquents que *pouvoir* et *devoir*. Au vu de cette propriété, ils ont moins de chance de rencontrer d'autres verbes et donc de se démarquer par des associations particulières. Notons également que *falloir* et *vouloir* ont tous deux la capacité d'omettre une complémentation verbale sous

forme de verbe à l'infinitif, ce qui peut également expliquer leur proximité par leur manque d'association avec des verbes.

Pouvoir, *falloir* et *vouloir* sont ainsi les verbes qui s'associent le plus aux verbes de *communication* et ceux groupés sous l'étiquette *psychologique*. Ces verbes ont des caractéristiques proches des modaux, « [e]n effet, comme ceux-ci, ils modalisent l'énoncé » (Gómez-Jordana Ferary & Anscombre 2015 : 7) et ils « introduisent aussi une représentation du monde effectuée par la parole, avec spécification du type de parole utilisé, de sa provenance et de la position du locuteur » (Anscombre 2015 : 103). Il semble ainsi que ces associations démontrent d'une double modalisation de l'énoncé importante dans les corpus que nous avons utilisés pour notre étude. Ceci pourrait être lié à la théorie de la politesse de Brown et Levinson (1987) : l'usage des verbes modaux adoucit ou met une distance sur des *paroles* et des *pensées* relatées dans des textes supposément objectifs¹⁸⁸.

Devoir est donc être le verbe qui se distingue le plus des autres modaux puisqu'il ne s'associe pas de manière aussi importante avec les verbes de *communication* ou verbes *psychologiques*. Les catégories qui composent son profil sémantique sont beaucoup plus disparates que pour les autres modaux, en particulier *falloir* et *vouloir* qui montrent une prédominance nette pour les verbes *psychologiques* ou de *communication*. Il se rapproche cependant de *pouvoir* dans son usage épistémique – en particulier dans le corpus de presse – et de *falloir* au présent de l'indicatif lorsque lui est conjugué au conditionnel présent dans les AFC présentées. Avec *vouloir*, il partage le fait d'attirer les verbes dénotant des actions plus concrètes que celles véhiculées par les verbes *dicendi* ou *putandi*.

Ainsi, concernant nos hypothèses de base au sujet des associations des modaux avec leur complémentation verbale, celles-ci ont pu être quelque peu nuancées. Certes, *pouvoir* et *devoir* se ressemblent dans leurs usages épistémiques mais les rassemblements dénotés dans cette partie concernent surtout *pouvoir*, *falloir* et *vouloir* et leurs associations avec les verbes de *communication* ou verbes *psychologiques*. Dans une étude future, il pourra être intéressant d'affiner ces catégories afin de distinguer de manière plus précise les caractéristiques individuelles ou communes de ces verbes modaux.

Comme nous le notions également pour les analyses concernant les sujets grammaticaux qui prennent la forme d'un nom, il semble également essentiel de prendre en compte l'ensemble de l'énoncé pour pouvoir interpréter le modal et lui attribuer une certaine lecture sémantique. Dans cette description globale, il faudrait également tenir compte du temps et du mode verbal du

¹⁸⁸ Ceci semblait être également le cas avec certains noms en position sujet, notamment les noms dénotant des normes ou des *a priori* pour *vouloir* ou encore des termes relatifs à des règlements pour *devoir*.

modal qui jouent un rôle déterminant dans les caractéristiques sémantiques qu'on lui attribue. Cependant, contrairement aux sujets nominaux, l'inférence semble ici moins nécessaire pour interpréter la valeur modale ou l'effet de sens véhiculé par le modal. Le verbe à l'infinitif se trouvant en position de complémentation verbale pour *pouvoir*, *devoir*, *falloir* et *vouloir* est ainsi plus révélatrice des caractéristiques sémantiques du modal que le sujet grammatical représenté par un nom.

7.3. Synthèse de l'analyse des cotextes des verbes modaux

Le but de cette analyse des cooccurrences spécifiques de *pouvoir*, *devoir*, *falloir* et *vouloir* était non seulement de confirmer des propriétés sémantiques qu'on leur octroyait jusqu'à présent dans les études à leur propos mais également de voir si des propriétés sémantiques encore inconnues pouvaient leur être attribuées. Une originalité non négligeable de cette recherche est également le fait que ces quatre verbes modaux n'ont jamais été comparés entre eux dans une étude quantitative. Nous résumons ici les principaux intérêts et résultats de notre recherche.

7.3.1. Intérêt de cette analyse

L'importance du co(n)texte des modaux pour en déterminer les caractéristiques sémantiques est attestée depuis longtemps mais n'avait jamais été étudiée autant en profondeur ni avec des outils statistiques de la linguistique de corpus. Ces réflexions existantes sont très vagues et ne permettent pas des généralisations¹⁸⁹.

Pour les sujets grammaticaux, il existe une certaine contradiction à propos du rôle qu'ils peuvent jouer dans la lecture sémantique du modal : il est admis que *vouloir* est un verbe à *partage*, car il partage le sujet de l'infinitif apparaissant dans la complémentation alors que *pouvoir* et *devoir* sont des verbes à *héritage*, qu'ils *héritent* ainsi du sujet du verbe qu'ils modalisent. Pourtant, un verbe à *partage* devient à *héritage* dès lors qu'il possède des sujets d'autres types que de nature *animé*, ce qui est le cas pour *vouloir*. De plus, des études admettent que certains noms en position sujet influencent l'interprétation de la valeur modale ou de l'effet

¹⁸⁹ Rappelons que c'est un des avantages principaux des méthodes quantitatives qui permettent, à juste titre, le repérage de faits généralisables : « Il est rare que les linguistes s'intéressent aux échantillons en tant que tels, plutôt qu'aux généralisations à partir des échantillons vers la quantité infinie de texte correspondant à la définition extensionnelle d'une (sous-)langue. Par exemple, un linguiste étudiant un modèle dans les 500 échantillons de texte du corpus Brown sera typiquement intéressé à tirer des conclusions sur l'anglais américain (écrit) dans son ensemble, et non seulement sur les textes spécifiques qui composent le corpus Brown » (Baroni & Evert 2008 : 777).

de sens de *pouvoir* et *devoir*, alors même qu'ils ne choisiraient pas leur sujet. Nous avons donc pris le parti d'étudier les sujets grammaticaux en ne tenant pas compte de cette distinction qui peut être contredite.

Concernant la complémentation verbale sous forme de verbe à l'infinitif, celle-ci distingue les modaux d'autres types de verbes. Elle pourrait ainsi avoir plus d'influence sur leur interprétation que le sujet grammatical.

C'est sur la base de ces remarques que nous avons donc mené notre recherche qui était tant corpus *based* que *driven* : *based* car nous nous basons sur les hypothèses qui formulent que le sujet grammatical ainsi que la complémentation verbale ont une influence sur les propriétés sémantiques des modaux ; *driven* car nous ne nous attendions pas à des résultats en particulier, hormis le fait que les usages grammaticalisés de ces modaux devaient se retrouver parmi les cooccurrences significatives des profils combinatoires.

Afin de mener au mieux cette étude, nous avons opté pour des classifications sémantiques des noms et des verbes pour analyser les résultats de manière plus optimale : en effet, ces groupements nous ont permis de faire ressortir des caractéristiques de manière plus prégnante que si nous avions dépouillé tous les cooccurents spécifiques un par un pour chacun de ces verbes modaux. Nous avons donc proposé une réflexion sur les diverses classifications existantes à ce propos pour finalement arrêter notre choix sur la typologie des noms de Salvadori et al. (2021) et le dictionnaires *Les Verbes français* de Dubois et Dubois-Charlier (1997).

Notre recherche nous permet aussi d'apporter des nuances aux propriétés combinatoires des modaux car nous avons fait varier plusieurs paramètres comme la forme du verbe (lemme, présent de l'indicatif et conditionnel présent) mais aussi le type de texte dans lesquels ils étaient usités (presse écrite nationale ou encyclopédie collaborative). Cela a donné l'occasion de mettre en avant des caractéristiques stables de ces verbes mais aussi de mettre en exergue les facteurs liés aux genres et champs génériques qui peuvent exercer une influence sur leurs particularités. Les méthodes usées sont également diversifiées pour pouvoir corroborer ou, du moins, préciser nos résultats. Les indices de spécificité des cooccurrences avec les sujets et les complémentations verbales ont donné l'occasion de révéler les propriétés individuelles de chacun de ces modaux, étayées avec des exemples de nos corpus. Les AFC et les analyses arborées ont ensuite mis en évidence leurs ressemblances et divergences. Cela a permis de révéler des propriétés qu'on ne leur avait pas encore attribuées jusqu'à présent et surtout de mettre sur un même plan ces modaux, qui n'avaient jamais été rapprochés dans une même étude. Les profils combinatoires, AFC et analyses arborées nous ont révélé leurs préférences

combinatoires mais c'est ensuite la lecture des exemples des corpus qui a permis d'étayer leurs propriétés sémantiques, ce qui rappelle la remarque de Leblanc (2015) qui précise qu'un retour au texte est nécessaire afin de pouvoir analyser au mieux les résultats purement quantitatifs (Leblanc 2015 : 53). C'est pourquoi nous insistons encore une fois sur le fait que ces analyses sont complémentaires mais non pas substituables à celles présentées avec l'indice de spécificité. Mentionnons également que ces types de représentation sont rares pour l'étude des cooccurrences, et nous espérons avoir pu montrer ici comment ils peuvent s'avérer utiles pour jeter un regard plus macroscopique sur des ressemblances et des divergences entre des unités lexicales comparables.

7.3.2. Les résultats à retenir

Plusieurs points sont à retenir de cette analyse. Nous rappelons ici ce que cette recherche a permis de révéler de nouveau à leur propos et les résultats qui ont confirmé certaines propriétés qu'on attribuait déjà dans les études jusqu'à présent à ces verbes modaux.

7.3.2.1. Les propriétés qui ont été confirmées

Les calculs statistiques ont permis d'attester des traits sémantiques attribués aux modaux. Ainsi, c'est la catégorie des noms *animés* qui ressort le plus souvent pour *vouloir*, peu importe sa forme conjuguée et le corpus investigué. Son NS d'*intention/volonté* se confirme ainsi par cette propriété combinatoire.

Les emplois grammaticalisés sont aussi ressortis de manière très fréquente pour les verbes modaux. Nous avons ainsi pu avoir de nombreux usages de *faut croire*, *faut imaginer*, *faut voir* ou encore *vouloir dire*.

Quant aux interprétations sémantiques de ces modaux, nous avons en effet pu constater que les effets de sens et valeurs modales de *pouvoir* sont difficiles à délimiter, ce qui avait déjà été soulevé par Le Querler (2004 : 653) et les sujets *animés* confèrent à *devoir* un effet de sens d'*obligation*.

7.3.2.2. Les propriétés inédites des verbes modaux

Cette étude a permis de mettre en avant que, pour interpréter sémantiquement les verbes modaux, il était nécessaire de prendre en compte l'ensemble du co(n)texte, en particulier pour

les cooccurrences liées aux sujets grammaticaux sous forme de nom qui pouvaient difficilement assurer une interprétation sémantique du modal. Pour les verbes à l'infinitif en position de complémentation verbale, l'interprétation était plus systématique. Ainsi, nos exemples ont démontré que *demander* confère une valeur modale déontique à *pouvoir*, *permettre* une valeur modale épistémique à *devoir*. Cependant, ce sont des mécanismes d'inférences qui nous ont permis de déclarer que les sujets représentant des instances d'autorité attribuent une valeur modale déontique aux modaux. Les profils combinatoires ont également démontré une variation selon le temps ou le mode verbal, voire selon le corpus, ce qui révèle une certaine volatilité du NS – ou du moins dans l'interprétation – des verbes modaux.

Nous allons désormais détailler les différentes propriétés à retenir pour chacun de ces verbes et également les ressemblances et divergences que ces résultats ont mis en lumière.

7.3.2.3. Devoir

Nous tenons à commencer notre rétrospective des résultats par *devoir* car, comme signalé à plusieurs reprises, c'est le modal qui ressemble le plus à son NS : en effet, dans les différents exemples que nous avons dépouillés, nous avons pu noter que ses valeurs modales et effets de sens se rattachent toujours à son NS de manière aisée et qu'il reproduit ainsi les propriétés qu'on lui attribue généralement dans la littérature.

Par rapport aux autres verbes, les divers profils combinatoires des sujets nominaux ou des complémentations verbales démontrent une répartition assez égalitaire des résultats. Ainsi, il ne se démarque pas en attirant une catégorie de noms ou de verbes en particulier dans sa valence.

Dans les divers effets de sens et valeurs modales relevées à l'aide des occurrences des corpus, nous avons pu noter beaucoup de valeurs modales déontiques mais aussi des énoncés dans lesquels il permettait d'émettre une *prédiction*. La lecture de *prédiction* est possible dans les cas où il y a mention d'une date dans l'énoncé, ce qui démontre une fois de plus l'importance de l'inférence dans l'interprétation sémantique des verbes modaux.

Dans les verbes qu'il attire en position de complémentation verbale, nous avons pu noter une préférence pour ceux qui dénotent des actions « concrètes » contrairement aux autres modaux qui préfèrent la compagnie des verbes de *communication* ou *psychologiques*. Ces attractions montrent une certaine inaptitude de *devoir* à se grammaticaliser. En effet, les autres modaux ont tendance à perdre leurs NS avec les verbes de *communication* et *psychologiques*. *Devoir*, lorsqu'il est attiré par ce type de verbes, garde son NS initial. Ceci rejoint ainsi nos réflexions

émises quelques lignes plus haut : *devoir* est le modal le plus fidèle à son NS et donc aux catégories traditionnelles de la modalité.

7.3.2.4. Pouvoir

Concernant *pouvoir*, nous avons pu confirmer ce que les études ont dit à son propos jusqu'à présent : les effets de sens sont nombreux pour ce verbe. Ceci peut être dû à ses profils combinatoires qui montrent une grande variété d'attractions de catégories de noms et de verbes. En ceci, nous pouvons rappeler la remarque de Bertels et Speelman (2012) au sujet de la variété des co(n)textes d'utilisation d'une unité lexicale :

Une unité lexicale monosémique apparaît dans des contextes plutôt homogènes sémantiquement, parce qu'elle se caractérise par des cooccurents qui appartiennent à des champs sémantiques plutôt similaires. En revanche, une unité lexicale polysémique se caractérise par des cooccurents plus hétérogènes sémantiquement, qui appartiennent à des champs sémantiques différents (Véronis 2003 ; Habert et al. 2004). Habert et al. (2004) avancent l'hypothèse qu'un mot à sens multiple "aurait des voisins moins proches entre eux qu'un mot univoque" (Habert et al. 2004 : 570). Les mots homonymiques, polysémiques et vagues auraient donc des cooccurents sémantiquement plus hétérogènes. (Bertels & Speelman 2012 : 149)

Au vu de ses usages très variés, nous pouvons nous demander si c'est véritablement le NS de *possibilité* qui transparait pour *pouvoir*. Rappelons que Gosselin (2010 : 449) propose un NS de *compatibilité* pour ce modal et Kronning (1996) celui de *problématicité*. Pourtant, ces NS ne semblent pas suffisants pour expliquer les usages de *pouvoir* qui sont moins modaux. Cependant, même en admettant un autre NS pour ce verbe, cela ne permet pas de rendre compte de ses nombreux usages autres que ceux dénotés par la littérature, que nous pouvons qualifier de grammaticalisés. En effet, nous avons pu noter que ce modal permet d'exprimer des généralités, surtout avec le pronom *on* et qu'il se grammaticalisait avec les verbes de *communication* et verbes *psychologiques*. Dans ces cas, il occupe un rôle de marqueur de focalisation, permettant de mettre en avant une *pensée* ou un *dire*, mais aussi de présenter des pensées alternatives ou encore des hypothèses. Ceci rejoint en quelque sorte la notion de « micro-contexte d'effet abstraitisant » soulevé par Larsen (1984) (in Fløttum et al. 2007 : 22-23) avec le pronom *on* mais aussi les remarques de Sitri (2013) dont ce type d'emploi pouvait être paraphrasé par « le procès a eu lieu et c'est notable » (Née et al. 2016 : 83).

7.3.2.5. Falloir

Falloir n'a pas beaucoup été analysé du point de vue de la modalité et les résultats à son propos permettent quelque peu d'explicitier pourquoi.

Il n'existe à notre connaissance pas de recherche sur le sujet exprimé dans la complétive. Notre étude a pu relever qu'il s'agissait dans la majorité des cas de sujets *animés* et que dans ces occurrences, *falloir* permettait ainsi d'exprimer un but à accomplir pour ces personnes, mais de manière quelque peu distancée, ce qui pourrait ainsi rejoindre les remarques de Brown et Levinson (1987) sur l'usage des formes impersonnelles pour ménager la face. Sa nature impersonnelle confère à ces associations une certaine prise de distance sur l'*obligation* portée sur les personnes ou les groupes sur lesquels porte la modalité de *falloir*.

Dans les AFC, ce verbe apparaît toujours au centre des deux axes, et, comme nous l'avons relevé à plusieurs reprises, cette position démontre un manque de spécificité par rapport aux paramètres étudiés. Ainsi, par rapport aux autres modaux, il ne semble pas se démarquer, ni sur les sujets de sa complétive, ni sur le verbe qui suit. Ce manque de différenciation peut s'expliquer selon deux facteurs que nous avons relevés. Le premier est que, par sa nature de verbe impersonnel, *falloir* est simplement moins sensible que les autres à son co(n)texte. Le second est que, dans les cooccurrences les plus spécifiques de *falloir*, nous retrouvons les verbes *psychologiques*. Avec ces verbes, *falloir* se grammaticalise et agit comme un marqueur de focalisation ou d'alternative. Ses usages relevant des catégories traditionnelles de la modalité sont finalement assez rares mais ceci peut être un biais de nos corpus.

7.3.2.6. Vouloir

Les profils combinatoires de *vouloir* ont, certes, confirmé son attraction importante avec les *animés* tout comme avec le verbe *dire*. Cependant, d'autres caractéristiques ont émergé de notre étude.

Ainsi, une part importante des sujets grammaticaux spécifiques à *vouloir* sont représentés par la catégorie des *objets cognitifs*. Tous les usages que *vouloir* possède avec ce type de noms sont grammaticalisés, dans le sens où *vouloir* n'exprime plus le NS d'*intention/volonté* qu'on lui attribue généralement. Dans ces associations, *vouloir* permet d'exprimer des *normes* ou encore des *suppositions*. Ce type d'effet de sens est plus fréquent dans le corpus encyclopédique que dans le corpus de presse, dans lequel *vouloir* est plus fidèle à son NS.

Les verbes de *communication* sont très importants en position de complémentation verbale pour *vouloir*. Ceci est en grande partie dû au verbe *dire* avec lequel le verbe se fige. Avec tous les autres verbes de *communication*, *vouloir* exprime des effets de sens plus traditionnellement associés à son NS.

Vouloir partage donc des traits communs avec tous les autres modaux étudiés ici : il ne se démarque pas tant par ses associations – comme c’est le cas pour *falloir* –, il permet des usages grammaticalisés avec *dire* – ce qu’on a pu observer pour *pouvoir* également – et exprime dans les autres cas son NS d’*intention/volonté* de manière très régulière – comme relevé pour *devoir*. *Vouloir* a également un comportement très stable dans ses profils combinatoires mais aussi dans les AFC et les analyses arborées présentées, et ce, quel que soit le temps ou le mode verbal auquel il est conjugué.

7.3.3. Similarités et disparités entre les verbes modaux

Un des premiers phénomènes que nous pouvons souligner ici est le fait que les verbes modaux partagent plus de points communs concernant leurs complémentations verbales que par rapport à leurs sujets grammaticaux. Les seuls verbes qui possèdent ici des ressemblances importantes concernant leurs sujets sont *pouvoir* et *devoir*. En effet, ces deux modaux présentent des profils combinatoires similaires dans l’étude des sujets. Les catégories les plus dominantes sont souvent les mêmes et c’est seulement dans le bas du classement des classes sémantiques les plus représentées que des distinctions se dessinent. *Pouvoir* et *devoir* se ressemblent particulièrement en tant que lemme et au conditionnel. Pour ce mode verbal, c’est donc l’expression de la valeur modale épistémique qui les rapproche. Dans WIKI, nous avons par exemple pu observer que ce mode permet aux deux verbes de véhiculer des *hypothèses*. Il est également important de souligner que dans les analyses arborées, ils se retrouvent systématiquement du même côté du nœud lorsqu’on les étudie en tant que lemmes, que ce soit pour les sujets ou les complémentations verbales, dans LM ou dans WIKI. Donc malgré toutes les différences que nous avons pu soulever le long de ce travail, il y a bien des rapprochements à noter pour ces deux modaux aux NS distincts ; l’usage de la valeur modale épistémique est suffisamment important pour ces deux verbes pour les rassembler systématiquement dans ce type de graphique.

Pouvoir et *falloir* partagent également des caractéristiques communes. La plus importante est leur capacité d’agir en tant que marqueur de focalisation sur des généralités avec les verbes dits

psychologiques, *falloir* par sa nature de verbe impersonnel et *pouvoir* lorsqu'il est précédé par le pronom *on*.

Falloir et *vouloir* se ressemblent dans le fait qu'ils sont plus homogènes que les deux autres modaux dans leurs profils combinatoires. Cette similarité peut être due au fait que ces deux verbes n'ont pas nécessairement besoin d'un verbe à l'infinitif pour porter un effet de sens et que, lorsque celui-ci est exprimé, il s'agit d'un certain type de verbe en particulier, à savoir les verbes *psychologiques* pour le premier et ceux de *communication* pour le second. Ce sont également les modaux les plus réguliers dans leurs comportements car leurs profils combinatoires et positions dans les AFC sont semblables quelle que soit la forme du verbe (en tant que lemme, au présent de l'indicatif ou au conditionnel présent) ou encore le corpus étudié. *Pouvoir* et *vouloir* apparaissent toujours à proximité l'un de l'autre dans les AFC. Ceci peut être dû au fait qu'ils présentent une ambivalence similaire : soit ils expriment les valeurs modales et effets de sens qui leur sont propres, ou alors, il se grammaticalisent avec les verbes de *communication* ou encore les sujets de type *objet*. En effet, avec ce type de sujet, *pouvoir* exprime des faits généraux, ce qui l'éloigne ainsi de son NS de *possibilité* traditionnellement admis.

Devoir et *vouloir* sont également analogues pour certaines de leurs propriétés combinatoires. Ainsi, ces deux verbes se rapprochent dans les AFC du côté des sujets *animés* mais aussi des verbes dénotant des actions « concrètes ». Par ces associations, ils expriment leurs NS respectifs, à savoir celui de la *nécessité* ou de l'*intention/volonté* qui peuvent ainsi être utilisés dans des co(n)textes similaires.

Devoir et *falloir* ont finalement peu de points communs alors qu'ils sont généralement considérés comme des prototypes de la valeur modale déontique. Leurs rapprochements sont surtout à relever dans les AFC qui rassemblent les usages au présent de l'indicatif et au conditionnel présent. Dans ces graphiques, la forme au conditionnel de *devoir* et celle au présent pour *falloir* sont adjacentes. Cela semble ainsi montrer que c'est majoritairement dans ses usages de valeur modale épistémique – et donc de marqueur d'incertitude avec une certaine prise de distance sur le *dit* – que *devoir* ressemble à *falloir*, qui lui, marque cette prise de distance par sa simple nature de verbe impersonnel.

Une des ressemblances majeures concerne cependant *pouvoir*, *falloir* et *vouloir*. Comme nous l'avons soulevé, ils attirent majoritairement des verbes de *communication* et/ou des verbes *psychologiques*. Toutefois, ce n'est pas uniquement le fait qu'ils attirent le même type de verbes qui les fait se ressembler mais surtout leur faculté à se grammaticaliser tous les trois avec ces catégories-là. Cette capacité à perdre leur NS les regroupe entre eux mais les éloigne surtout de

devoir. Ceci a été confirmé tant dans leurs profils combinatoires que dans les diverses AFC montrées dans ce travail.

Enfin, notons également que les modaux étudiés ici se distinguent davantage dans WIKI que dans LM, ce qui montre des usages moins différenciés entre ces verbes dans la presse que dans le discours encyclopédique.

7.3.4. Limites de cette étude

Même si cette recherche permet de révéler des faits nouveaux sur les verbes modaux, elle présente aussi quelques limites dont nous avons conscience et qui nous semblent importantes de rappeler ici.

Ainsi, les résultats analysés concernent toujours les indices de spécificité les plus importants. Les conclusions que nous tirons donc pour les différentes propriétés sémantiques des verbes modaux concernent seulement leurs usages les plus fréquents. Il n'est donc pas possible de prétendre ici à une exhaustivité des propriétés sémantiques des verbes modaux analysés dans cette thèse.

Les exemples que nous avons dépouillés pour illustrer les différents usages des modaux ne sont pas exhaustifs non plus. Nous avons à chaque fois étudié seulement un certain échantillon des occurrences des modaux avec tel sujet ou telle complémentation verbale. Nous n'avons pas pu quantifier le nombre de fois où tel effet de sens ou valeur modale survenait dans la totalité des cooccurrences spécifiques. Il se pourrait ainsi que les usages que nous avons relevés soient biaisés par les exemples analysés dans le cadre de notre échantillon.

Enfin, le fait que nous ayons été la seule personne responsable des classements des divers verbes ou noms dans les catégories de Dubois et Dubois-Charlier (1997) et Salvadori et al. (2021) ne garantit en aucun cas une précision des résultats qui pourraient être nuancés avec une deuxième annotatrice.

7.3.5. Perspectives de recherche

Par la nature de notre recherche, nous avons proposé ici une sorte de « catalogue » d'exemples et de cas qui pourraient être intéressants à approfondir. Par exemple, le facteur du temps ou du mode verbal pourrait encore être étayé en ajoutant notamment le futur ou encore l'imparfait dans les paramètres de recherche. L'imparfait, par exemple, peut exprimer la *contreactualité* et l'*hypothèse* (Heenen 2016) et donc la valeur modale épistémique, qui serait ainsi intéressante

à mettre en parallèle avec le conditionnel. Quant au futur, il permet de véhiculer une valeur modale épistémique mais aussi une « injonction déguisée » (Ricci 2017 : 33) et donc une valeur modale déontique.

Il serait évidemment intéressant d'étudier ces modaux dans d'autres types de textes – des textes oraux notamment pour analyser de manière plus détaillée la sémantique des verbes modaux dans le cadre des actes de langage notamment – mais aussi dans un corpus de référence. Un corpus de ce type, par sa grande taille, son équilibre et la diversification des genres textuels représentés (Siepmann et al. 2016) permettrait de poser de manière plus affirmée les propriétés sémantiques des verbes modaux.

D'autres éléments du cotexte pourraient également être pris en compte afin d'affiner nos remarques sur les propriétés combinatoires des verbes modaux. Les adverbes, notamment, qui permettent de nuancer l'effet de sens ou la valeur modale véhiculée¹⁹⁰, mais aussi les sujets grammaticaux autres que les noms, notamment les pronoms personnels mais aussi les noms propres. Nous avons par exemple pu soulever que le pronom *on* conférait à *pouvoir* une certaine capacité à exprimer des faits généralisants. Cette piste serait ainsi intéressante à approfondir.

Dans cette thèse, nous avons séparé les analyses sur les sujets et celles sur la complémentation verbale. Il pourrait être intéressant de s'intéresser à des cooccurrences plus vastes, par exemple, étudier quelle catégorie de verbe en particulier est attirée par les *animés*, suivis par *devoir* et de prendre ainsi en considération une chaîne de mots plus importante pour un examen encore plus détaillé de certains emplois.

Il serait évidemment pertinent de se focaliser sur les usages plus grammaticalisés de ces verbes afin de comprendre dans quels cotextes, c'est-à-dire avec quels verbes et sujets en particulier ce type de phénomène est possible.

¹⁹⁰ A ce propos, nous pouvons par exemple citer la thèse de Hütsch (2020) qui s'est focalisée sur l'association de *pouvoir* et *devoir* avec des adverbes modaux.

8. Conclusions et perspectives générales de la thèse

Lors d'une conférence en octobre 2023 à Kalamata en Grèce, une chercheuse est venue vers nous à la fin de notre présentation et nous a dit : « Je pensais qu'il n'y avait plus rien à dire sur les verbes modaux en français, votre présentation m'a prouvé le contraire : le champ des possibles est encore énorme ! ». Nous espérons que cette thèse a pu démontrer cette même idée. Dans cette partie conclusive, nous allons revenir sur les objectifs que nous nous étions fixés (8.1.) et nous reprendrons les principaux résultats de notre analyse (8.2.). Ces points nous permettront de mettre en exergue les apports de la thèse (8.3.), avant de proposer plusieurs perspectives de recherches que nous – ou d'autres chercheuses intéressées par la thématique – pourrions mener (8.4.).

8.1. Rappel des objectifs de cette thèse

Au début de ce travail, nous avons posé plusieurs questions de recherche auxquelles nous avons répondu à travers les différentes analyses menées ici. Nous allons brièvement les repasser en revue.

Un des premiers points que nous avons soulevés était la question même d'étudier les caractéristiques sémantiques des verbes modaux. Nous avons rappelé que plusieurs linguistes considèrent ces verbes comme de simples (semi-)auxiliaires notamment à cause de leur valence particulière, qui les oblige à précéder un verbe à l'infinitif. En outre, nous avons mentionné que les modaux étaient davantage considérés dans des études sur l'énonciation et donc dans un cadre relevant davantage de la pragmatique que de la sémantique. Par les différentes propriétés sémantiques relevées dans la littérature et présentées dans le chapitre 4 de cette thèse, nous avons pu démontrer que les verbes modaux permettent d'exprimer des *effets de sens* et des *valeurs modales* à partir de leurs *noyaux sémantiques* (NS). Il semble dès lors impossible de les mettre au même niveau que les auxiliaires *être*, *avoir* ou encore *aller* car ils ont des caractéristiques sémantiques importantes. Cependant, nous avons aussi très vite admis qu'il était impossible de se détacher complètement des théories pragmatiques car les mécanismes d'inférence sont parfois – pour ne pas dire *souvent* – nécessaires pour interpréter l'effet de sens ou la valeur modale d'un de ces verbes.

A partir de ce constat, il a également été relevé que certains usages de ces verbes modaux étaient similaires. Par exemple, *pouvoir*, *devoir*, *falloir* et *vouloir* peuvent tous – à leur échelle – exprimer les valeurs modales épistémiques et déontiques. Ils partagent donc des points

sémantiques communs alors que leurs NS sont distincts – hormis peut-être ceux de *devoir* et *falloir* à qui on attribue un NS de *nécessité* pour le premier et de *nécessité* ou de *contrainte* pour le second. Notre postulat était que même si ces verbes se ressemblent – car ils peuvent véhiculer des valeurs modales identiques – ils doivent présenter des différences notables. Mais, si ce ne sont pas les catégories traditionnelles de la modalité qui arrivent à faire émerger leurs divergences, qu'est-ce qui permet de le faire ? C'est pourquoi nous nous sommes concentrée ici sur leurs profils combinatoires et donc leurs effets de sens et valeurs modales en *usage* en voulant tester diverses hypothèses.

8.2. Principaux résultats de cette thèse

Les recherches que nous avons menées à l'aide de méthodes empruntées tant à la tradition anglo-saxonne de la linguistique de corpus qu'à la vision plus francophone de la textométrie (présentées dans le chapitre 6 de ce travail), nous ont permis non seulement de vérifier les diverses hypothèses que nous avons émises au sujet des ressemblances et divergences des verbes modaux, mais aussi de relever des propriétés encore méconnues à leur propos. Ce sont ces résultats que nous rappelons dans cette section.

8.2.1. Vérification des hypothèses de recherche

Nous allons évoquer dans cette section les diverses hypothèses que nous avons proposées dans le chapitre 5 de cette thèse et les résultats qui permettent de les confirmer ou de les infirmer.

- i) Les phénomènes de grammaticalisation des modaux peuvent être vérifiés avec leurs associations privilégiées, mesurées à l'aide des outils statistiques de la linguistique de corpus.

Oui, les locutions grammaticalisées des verbes modaux – et déjà présentées dans la littérature à leur sujet – sont percevables à travers les cooccurrences spécifiques dans les deux corpus étudiés. Ainsi, *vouloir dire*, *faut croire/voir* et *je dois dire que* apparaissent parmi les usages les plus spécifiques aux verbes modaux. Pour les locutions intégrant *falloir* et *devoir*, cela peut paraître quelque peu étonnant car ces tournures sont davantage propres à la langue orale et les deux corpus investigués ici représentent une utilisation de la langue écrite, ce qui montre l'importance de ce phénomène et de ces associations.

- ii) Devoir et falloir, par leurs NS proches, voire identiques, auront des usages similaires dans les corpus étudiés.

Devoir et *falloir* ne partagent pas autant de traits communs que ne le laisse suggérer leurs NS. Ainsi, les divers profils combinatoires relevés, tout comme leurs positions sur les AFC et les analyses arborées étudiées, montrent que ces deux modaux ont des usages distincts. Cela est en grande partie dû au fait que *falloir* s'associe de manière importante aux verbes classés comme *psychologiques* et ces cooccurrences démontrent un usage grammaticalisé de *falloir*. Ainsi, même si leurs NS sont en théorie semblables, *falloir* est beaucoup plus sujet aux phénomènes de grammaticalisation, du moins, dans les corpus que nous avons utilisés pour cette thèse.

- iii) Même si pouvoir et devoir sont souvent comparés pour leurs usages de la valeur modale épistémique, leurs NS respectifs sont trop distincts pour que leurs cotextes d'utilisation soient comparables.

Cette hypothèse s'est avérée dans notre étude. *Pouvoir* et *devoir* présentent des profils combinatoires similaires, en particulier au conditionnel présent qui, selon nos exemples, leur confère une valeur modale épistémique. Ceci est principalement le cas dans le corpus journalistique LM dans lequel la locutrice émet davantage de doutes que dans le corpus encyclopédique WIKI. Ils sont également plus proches dans leurs combinaisons avec les sujets nominaux qu'avec les verbes en position de complémentation verbale. Une hypothèse qu'il serait ainsi intéressante à vérifier est le fait que leurs NS semble conditionner le choix de la complétive mais non pas du sujet grammatical.

- iv) Vouloir est un verbe modal sémantiquement plus riche que les autres. Il peut ainsi présenter un comportement distinct des trois autres modaux que nous étudions ici.

Les diverses analyses de cette thèse montrent que *vouloir* partage des points communs avec les trois autres modaux étudiés. Comme *falloir*, il manque d'association avec des catégories sémantiques particulières, ce qui démontre une certaine banalité dans son comportement. De plus, *vouloir* présente une attraction significative avec les verbes de communication, tout comme *pouvoir*, et avec des verbes dénotant des actions plus concrètes, ce qui le rapproche de *devoir*. Il ne se distingue de ce fait pas nettement de ces trois modaux mais se situe plutôt à un croisement de leurs propriétés, croisement qui est ressorti notamment dans les diverses AFC que nous avons analysées dans cette thèse.

- v) Les associations privilégiées de ces verbes modaux permettent de confirmer ou alors de proposer un NS plus précis pour chacun d'entre eux.

Nos analyses ont plutôt révélé que les verbes modaux étudiés ici avaient tendance à se grammaticaliser et donc à moins souvent véhiculer des effets de sens ou des valeurs modales rattachables à leurs NS, en particulier pour *pouvoir* et *falloir*, ainsi que, dans une certaine mesure, pour *vouloir* mais seulement dans son association avec le verbe *dire*. *Devoir*, quant à lui, semble être le modal le plus fidèle au NS qu'on lui attribue généralement, soit celui de *nécessité*.

8.2.2. Retour sur quelques résultats importants de cette thèse

Nous allons ici rappeler quelques éléments importants de notre analyse, déjà détaillés dans la *synthèse de l'analyse des cotextes des verbes modaux* mais nous les mentionnons ici brièvement afin de revenir ensuite sur les apports de cette thèse.

8.2.2.1. Pouvoir est le plus flou des verbes modaux

Tout au long de l'analyse, nous avons pu relever que *pouvoir* était un verbe modal relativement « volatile », car les contours de son interprétation sémantique sont flous. Ainsi, contrairement aux autres verbes modaux étudiés ici, une cooccurrence significative avec une catégorie de noms en position sujet ou une catégorie de verbe en position de complémentation verbale ne permet pas d'interpréter de manière catégorique l'effet de sens ou la valeur modale véhiculée. Les exemples étudiés ne démontrent pas non plus un usage marquant de son NS de *possibilité*. En effet, *pouvoir* permet d'exprimer des généralités – notamment lorsqu'il est usité pour exprimer des valeurs modales descriptives avec des noms de type *objet* en position sujet – mais aussi lorsqu'il est suivi par un verbe de *communication* ou *psychologique* et précédé par le pronom *on*. Plus précisément, il permet de porter l'attention de la lectrice ou de l'interlocutrice sur ce fait général lorsqu'il est précédé par *on* et suivi d'un verbe de *communication* ou d'un verbe *psychologique*. Il se grammaticalise ainsi dans ses cooccurrences les plus significatives.

8.2.2.2. Devoir : le plus modal des modaux ?

Devoir est le verbe modal étudié ici qui est le plus facilement interprétable en usage. Les effets de sens et les valeurs modales qu'il véhicule dans les deux corpus investigués relèvent toujours

des notions sémantiques qu'on lui attribue traditionnellement dans la littérature. Dans tous les exemples recueillis, il a toujours été possible de rattacher l'interprétation de l'énoncé à son NS. Ceci peut être dû au fait qu'il est plus fréquemment associé à des verbes dénotant des actions concrètes, autres que la *parole* ou la *pensée*. En effet, pour les autres modaux, il semble que les phénomènes de grammaticalisation surgissent lorsqu'ils s'associent à des verbes de *communication* ou verbes *psychologiques*

8.2.2.3. Falloir : un marqueur de focalisation plutôt qu'un verbe modal ?

Alors que dans le cadre des études sur les marqueurs de la modalité *falloir* est la plupart du temps rangé aux côtés de *devoir* en tant que verbe dénotant une valeur modale déontique, nos analyses ont démontré que ce n'était pas son usage le plus fréquent.

En effet, il est, la plupart du temps, attiré par les verbes *psychologiques* et, avec cette catégorie, il agit en tant que marqueur de focalisation. Cet usage particulier l'éloigne des autres modaux étudiés ici, en particulier lorsque nous observons sa position dans les AFC. En effet, il se situe systématiquement au centre des deux axes. Nous pouvons ainsi nous demander si c'est sa nature syntaxique particulière de verbe impersonnel qui le différencie autant en usage. Est-ce que les verbes impersonnels sont moins marqués sémantiquement que les verbes à sujet et est-ce donc pour cela qu'il n'y pas de distinction marquée quant à son cotexte d'utilisation préféré ?

Une autre analyse possible relative à la position centrale des deux axes pour *falloir* peut être en rapport avec le NS de ce modal. En effet, celui-ci n'est peut-être pas aussi défini ou marqué que celui des autres verbes modaux, ce qui fait que les effets de sens et les valeurs modales ne correspondent ainsi pas aux connaissances que nous avons de ce modal. C'est pour cette raison qu'il ne se singularise pas avec ses associations spécifiques car son NS n'est ainsi peut-être pas aussi fixé que celui des autres verbes modaux étudiés ici. Ce qui peut également appuyer cette interprétation est le fait que, dans les usages que nous avons relevés, le verbe est soit grammaticalisé (avec les verbes *psychologiques* notamment), soit il exprime des concepts de *nécessité* avec une prise de distance relativement importante, due à sa nature de verbe impersonnel. Il n'y a ainsi pas une expression de son NS de manière aussi marquée que pour les autres modaux.

8.2.2.4. Vouloir : entre figement et expression de son NS

Dans la littérature, *vouloir* a été mis à distance des autres modaux – et même parfois déconsidéré en tant que modal – par sa nature sémantique plus riche que *pouvoir*, *devoir* et *falloir*.

Dans notre recherche, ces propriétés liées à son NS d'*intention/volonté* ont été confirmées par le fait qu'il attire majoritairement des sujets de type *animés*. Cependant, il cooccure également avec des noms de type *objets cognitifs* en position de sujet. Avec ce type de nom, nous avons pu démontrer que *vouloir* se grammaticalise pour exprimer des *règles*, des *normes* ou encore des *suppositions*. Ces usages sont très éloignés de son NS vu qu'on ne peut pas y lire une *intention* ou une *volonté*.

Les autres cas dans lesquels *vouloir* perd de son NS sont ses associations avec le verbe *dire* qui représente une des cooccurrences significatives les plus importantes. Ceci est le cas tant au présent de l'indicatif qu'au conditionnel présent et dans les deux corpus étudiés. Toutefois, c'est bien le seul verbe avec lequel il perd son NS alors que la grammaticalisation de *pouvoir* et *falloir* est plus générale avec toute une catégorie de verbes de *communication* ou de verbes *psychologiques*.

Ainsi pour *vouloir*, nous avons deux usages très distincts : ceux dans lesquels il exprime sans aucun doute son NS et les autres dans lesquels il se grammaticalise complètement.

8.2.2.5. Les ressemblances et divergences notables entre ces quatre verbes modaux

Les points communs que nous avons relevés entre les verbes modaux concernent surtout les verbes en position de complémentation verbale.

Les ressemblances les plus frappantes¹⁹¹ concernent surtout *pouvoir*, *falloir* et *vouloir* qui ont tendance à se grammaticaliser avec les verbes de *communication* et/ou avec les verbes *psychologiques*. Ce regroupement n'avait pas encore été relevé dans les études à leur sujet. Ce sont plus précisément *pouvoir* et *falloir* qui ont des usages similaires en tant que marqueurs de distanciation ou de focalisation sur le dire.

Notons également que les verbes modaux étudiés ici présentent des similarités ou divergences plus importantes selon le type de texte. Ainsi, dans WIKI, si on observe les sujets des verbes, il est intéressant de noter que *pouvoir*, *devoir* et *vouloir* permettent d'exprimer des généralités (avec la valeur modale descriptive pour *pouvoir* et les *objets*), des *normes* ou des *suppositions* (pour *vouloir* et les *objets cognitifs*) ou encore des *marches à suivre* (lorsque *devoir* est précédé

¹⁹¹ Les autres similarités sont à retrouver dans la *synthèse de l'analyse des cotextes des verbes modaux*.

d'un sujet *animé*). L'usage est donc ici différencié des usages que nous relevions dans le corpus de presse. Nous pouvons ainsi admettre que le type de texte peut avoir ici une influence sur les interprétations sémantiques des modaux.

8.2.3. Quelques remarques générales sur l'analyse

Notre analyse des caractéristiques sémantiques des verbes modaux a révélé que les effets de sens, les valeurs modales, voire les NS, traditionnellement attribués à ces verbes ne sont pas les plus fréquents observés en usage. En effet, notre recherche a mis en évidence des cooccurrences spécifiques particulièrement significatives pour ces verbes, révélant des emplois encore peu explorés dans les études antérieures. La méthodologie que nous avons adoptée, visant à mesurer initialement leurs indices de spécificité, a trouvé sa pleine justification lors de l'observation des associations significatives en usage, permettant ainsi de mettre en lumière des exemples jusqu'alors peu, voire pas, documentés dans les recherches à leur propos. Il convient de souligner que nous avons ensuite comparé ces différents verbes modaux à l'aide des méthodes d'AFC et d'analyse arborée, ce qui a permis de mettre en évidence visuellement leurs similarités et leurs différences.

Nous avons pu noter que la plupart des interprétations proposées dans l'analyse n'étaient pas seulement le fruit d'une cooccurrence avec un certain type de sujet ou de complémentation verbale et que c'est généralement le co(n)texte élargi, et donc l'inférence que nous en tirons, qui permet de comprendre le sens – ou la fonction – du verbe modal. En ceci, nous pouvons rappeler la réflexion de Zufferey et Moeschler (2012) :

En regardant une carte géographique, il est possible de suivre la ligne de frontière entre deux pays. En revanche, en voyageant, par exemple entre Genève et Paris en TGV, il n'est pas évident de savoir exactement quand on a passé la frontière. De la même manière, si nous pouvons nous donner des repères permettant de situer les questions de sens dans les domaines de la sémantique ou de la pragmatique, il est difficile de savoir exactement *quand* nous avons franchi la frontière. Par ailleurs, le domaine du sens est un peu comme le territoire de certains États d'Europe au cours de l'histoire : un jour il appartient à un pays, un autre à son voisin (comme par exemple l'Alsace ou la Pologne au cours des xix^e et xx^e siècles). (Zufferey & Moeschler 2012 : 179)

Tout au long de ce travail, nous avons cherché à étudier les caractéristiques sémantiques des verbes modaux, mais cette étude ne peut se faire sans les théories pragmatiques, notamment celles liées à l'inférence. De plus, il s'avère que les verbes modaux adoptent des fonctions autres qu'uniquement l'expression de certains effets de sens ou de certaines valeurs modales avec

leurs nombreux usages grammaticalisés. Ils permettent, par exemple, de nous focaliser plus précisément sur ce qui est dit lorsqu'on observe les usages de *pouvoir* et *falloir*. En ceci, nous remarquons ainsi qu'ils occupent des fonctions qui vont au-delà d'une catégorisation sémantique pure.

Avec ces résultats, nous pouvons nous demander à quel point il est nécessaire de parler de caractéristiques sémantiques des verbes modaux vu que nous avons pu démontrer que les cooccurrences les plus spécifiques pour *pouvoir*, *falloir* et *vouloir* sont des phénomènes de grammaticalisation et/ou de désémantisation.

8.3. Apports de la thèse

Nous proposons ici une réflexion plus générale sur l'apport de cette thèse, en nous concentrant, non pas sur la plus-value de notre analyse déjà présentée dans la section précédente, mais sur une contribution plus globale, portant sur les raisonnements que nous avons apportés dans le traitement même des verbes modaux avec les outils de la linguistique de corpus.

8.3.1. Implications des réflexions théoriques sur les verbes modaux

La thèse que nous avons proposée n'est pas une thèse sur la modalité mais sur les verbes modaux. Notre objectif était, en partie, de démontrer que les catégories modales généralement attribuées à ces verbes ne permettent pas de révéler pleinement leurs singularités : en effet, les travaux sur la modalité mettent surtout en avant les ressemblances entre les marqueurs modaux véhiculant les mêmes valeurs modales. Nous avons donc voulu définir en quoi ces verbes modaux se ressemblent ou divergent au-delà de ces catégories prédéfinies.

Pour cela, nous n'avons pas seulement étudié *pouvoir* et *devoir* – comme c'est traditionnellement le cas dans les études sur les verbes modaux en français – mais nous les avons insérés dans un paradigme plus large aux côtés de *falloir* et *vouloir*, deux verbes modaux rarement inclus dans les recherches sur les verbes modaux. Comparer ces quatre verbes a permis de faire ressortir des propriétés individuelles de chacun d'entre eux. En effet, si nous avions étudié seulement *pouvoir* par exemple, nous aurions pu montrer comment ce verbe fonctionne mais pas comment il se distingue dans ce fonctionnement même et donc en quoi il se particularise.

Au niveau de la terminologie, nous avons également apporté notre réflexion à propos des diverses notions qui gravitent autour de ces verbes modaux pour définir leurs propriétés

sémantiques. Nous avons ainsi différencié les notions d'*effet de sens*, de *valeur modale*, de *noyau sémantique* ainsi que défini ce que nous entendions par *phénomène de grammaticalisation*.

Dans l'état de l'art, il a été rappelé que ces verbes se ressemblent, car ils sont tous polysémiques et permettent ainsi d'exprimer plusieurs effets de sens et valeurs modales. Cependant, l'importance accordée à ces diverses notions n'est pas la même selon le modal étudié. C'est pourquoi nous avons individualisé le parcours réflexif pour chaque modal analysé ici, tout en faisant en sorte que les réflexions à leur propos puissent être comparées.

8.3.2. Implications des méthodes quantitatives pour l'étude des verbes modaux

Comme nous le rappelions au début de cette thèse, les propriétés sémantiques des verbes modaux en français ont rarement été étudiées à l'aide des méthodes outillées de la linguistique de corpus. Nous avons proposé de les aborder ici, non pas en partant des catégories traditionnelles de la modalité, mais en analysant les verbes modaux en usage, et plus précisément leurs cotextes d'apparition privilégiés.

Nous nous sommes ainsi attardée sur le cotexte immédiat de ces verbes, à savoir leur sujet grammatical (ou sémantique pour *falloir*) sous forme de nom ainsi que le verbe modalisé dans la complémentation verbale. Les recherches sur les verbes modaux avaient déjà analysé cet entourage pour déterminer les effets de sens et valeurs modales de ces verbes modaux sans toutefois faire appel à des méthodes quantitatives pour en tirer des généralisations.

Dans notre recherche, nous avons varié les méthodes afin de non seulement calculer les cooccurrences spécifiques avec l'indice de spécificité (Lafon 1980) de chacun de ces verbes et donc d'en tirer des profils combinatoires, mais aussi d'analyser les rapprochements et éloignements de ces formes à l'aide d'AFC et d'analyses arborées. Ces méthodes sont tant empruntées aux traditions anglophones de la linguistique de corpus qu'aux approches francophones de la textométrie. Mêler ces deux pratiques dans une étude des caractéristiques sémantiques est peu commun, et pourtant, nous avons pu voir comment elles se complètent afin de cerner, non seulement les propriétés individuelles d'unités lexicales, mais aussi de mesurer les similarités et disparités de termes proches. Mentionnons également qu'utiliser l'AFC et l'analyse arborée en tant que telles pour l'étude des cooccurrences est une démarche peu usitée encore et nous espérons que cette thèse a su démontrer en quoi elle pouvait être intéressante à mobiliser dans un cadre de recherche tel que celui adopté dans cette thèse.

En outre, cette analyse des verbes modaux a été menée dans deux types de textes différents qui nous ont permis d'apporter des nuances sur leurs comportements. Ces réflexions ont aussi été étayées en faisant varier le temps et le mode verbal des modaux afin de voir si les propriétés différaient selon ces critères.

Somme toute, notre méthodologie tant *corpus based* que *driven* a permis d'investiguer des pistes encore rarement considérées, notamment les ressemblances et dissemblances des verbes modaux par rapport à leurs associations avec les verbes de *communication* et les verbes *psychologiques*.

8.4. Ouvertures et perspectives de recherche

Comme toute recherche, cette thèse ouvre la voie à diverses perspectives et domaines d'exploration. Notre analyse approfondie des verbes modaux constitue une base solide pour de futures recherches et développements théoriques. Nous identifions et discutons ici quelques-unes des perspectives possibles issues de notre étude.

8.4.1. Étude d'autres facteurs pour les caractéristiques sémantiques des verbes modaux

Notre étude a traité de l'influence du sujet grammatical sous forme de nom et de la complémentation verbale sous forme de verbe à l'infinitif afin de faire ressortir des caractéristiques sémantiques des verbes modaux. Cependant, tout au long de notre analyse, nous avons pu relever que d'autres facteurs jouent un rôle dans l'interprétation sémantique des verbes modaux. Ces facteurs pourraient ainsi également faire l'objet d'études plus approfondies.

C'est le cas notamment du temps ou du mode verbal. Dans notre thèse, nous avons pu relever à l'aide de nombreux exemples que pour tous ces verbes, le mode du conditionnel leur confère des valeurs modales épistémiques, soulignant ainsi l'importance de ce mode dans l'interprétation sémantique des verbes modaux étudiés. Concernant le présent de l'indicatif, il permet à *devoir* d'exprimer la plupart du temps la valeur modale déontique. Pour les autres verbes, il semble que c'est ce temps-là qui est propice aux phénomènes de grammaticalisation. Il pourrait ainsi être pertinent d'étudier l'influence d'autres temps et/ou modes verbaux sur les interprétations sémantiques des verbes modaux.

Nous avons aussi relevé le fait que pour *pouvoir* en particulier, le pronom *on* permet des interprétations singulières de l'ordre de la généralisation sur la *pensée* ou le *dire*. Il serait ainsi

judicieux d'analyser de plus près les autres types de sujet et la façon dont ils influencent les interprétations sémantiques des verbes modaux.

Enfin, le poids du type de texte sur la sémantique des verbes modaux n'est pas à négliger. Nous avons pu constater que la valeur modale épistémique était davantage de mise dans le discours de presse alors que le discours encyclopédique était plus propice aux phénomènes de généralisations ou d'explications de normes et de règles à l'aide des verbes modaux. Il serait ainsi pertinent de mettre en contraste ces observations en étudiant l'usage des verbes modaux dans d'autres types de textes, notamment oraux.

8.4.2. Étude de la grammaticalisation des verbes modaux

Que ce soit dans la littérature ou dans l'analyse menée dans cette thèse, il a été soulevé que les phénomènes de grammaticalisation et/ou de désémantisation sont importants pour la compréhension du fonctionnement des verbes modaux. Creuser cette piste permettrait ainsi d'approfondir nos connaissances à leur propos. Cela pourrait s'effectuer d'un point de vue diachronique – pour mesurer à quel moment et de quelle façon le phénomène a surgi – en examinant de plus près les cooccurrences qui semblent se grammaticaliser. Nous pourrions ainsi observer plus en détail le type de sujet qui apparaît, par exemple, pour les locutions *vouloir dire* afin d'affiner la compréhension que nous avons de cette association ou regarder de plus près la catégorie des verbes de *communication* et des verbes *psychologiques* avec *pouvoir* afin de voir quels verbes exactement permettent une grammaticalisation de ce modal.

8.4.3. Annotations dans les corpus

Une dernière réflexion concernant les perspectives de recherche que nous proposons ici est davantage d'ordre méthodologique. En effet, il pourrait être envisageable d'utiliser les résultats exposés ici afin d'annoter les verbes modaux dans des corpus et ainsi d'étudier plus en détail le comportement des valeurs modales en usage. Nous pourrions ainsi étudier les associations privilégiées de tous les marqueurs que nous avons classés comme étant épistémiques par exemple. Toutefois, rappelons que définir de manière catégorique les valeurs modales de ces verbes est une difficulté non-négligeable car les contours des catégories modales sont floues, notamment à cause de leur polysémie, comme l'ont relevé Montrichard et al. (2022). Cependant, cette annotation peut justement être facilitée à l'aide des connaissances dont nous disposons désormais sur l'expression de certaines valeurs plutôt que d'autres en fonction de

l'entourage des verbes modaux. Ainsi, nous pourrions annoter toutes les constructions qui favorisent une interprétation épistémique, déontique, descriptive ou encore boulique. Ces annotations seraient, en revanche, seulement semi-automatiques. En effet, nous pouvons envisager d'extraire, par exemple, toutes les occurrences de *on + pouvoir au présent + un verbe de communication ou un verbe psychologique* car nous savons désormais que ce cotexte favorise une grammaticalisation de *pouvoir*. Un tri manuel serait tout de même nécessaire afin de vérifier si certains de ces usages n'expriment pas d'autres effets de sens.

Ces étapes pourraient aussi être facilitées si nous annotons directement dans les corpus les noms et les verbes avec les catégorisations de Salvadori et al. (2021) et Dubois et Dubois-Charlier (1997) que nous avons utilisées dans cette thèse. Dans notre recherche, nous avons trié, après l'extraction des résultats, les verbes et les noms selon les groupes définis dans les typologies susmentionnées. Il pourrait ainsi être intéressant de pouvoir mesurer une attirance générale d'un modal avec un groupe de noms et de verbes et non pas de mesurer parmi ses cooccurents spécifiques, lesquels appartiennent à quelle catégorie.

Ces étapes d'annotations pourraient également être facilitées si davantage de corpus étaient dotés d'annotations syntaxiques. Ainsi, il serait possible de savoir si les noms et les verbes qui cooccurrent avec les modaux dans l'empan déterminé font partie de leur valence. Nous pourrions, dans ces cas-là, parler de *motifs* dans lesquels les verbes modaux s'insèrent au sens de Legallois (2012) qui explique qu'un motif est « une unité multidimensionnelle, c'est-à-dire constituée à la fois d'associations lexicales ou grammaticales, d'appariements entre forme et sens ou fonction pragmatique / discursive, entre forme et fonction grammaticale » (Legallois 2012 : 45). Un motif, une fois identifié, est ainsi « une unité linguistique caractéristique d'un corpus (par exemple d'un genre, d'un auteur, etc.) » (Diwersy & Legallois 2014 : 1). Le repérage de ces motifs nous permettrait ainsi de présenter le fonctionnement des verbes modaux selon le sujet, le temps verbal, la complémentation verbale et le texte dans lesquels ils figurent.

8.5. Pour conclure...

Pour conclure cette étude des verbes modaux, nous pouvons rappeler que « [w]ords do not have meanings; meanings have words » (Williams 2003 : 96). Cette perspective souligne la nature dynamique et co(n)textuelle des significations linguistiques, un thème central de notre recherche. Nous avons en effet exploré comment les verbes modaux interagissent avec leurs co(n)textes, révélant ainsi leur rôle crucial dans la construction de la signification. Ils

représentent les mécanismes par lesquels les locutrices expriment des nuances de possibilité, de nécessité et de volonté. Notre analyse a montré que les verbes modaux se grammaticalisent, perdant parfois leur sens original pour adopter des fonctions nouvelles et co(n)textuelles. Cette transformation démontre la flexibilité et l'adaptabilité du langage, toujours en évolution pour répondre aux besoins communicatifs des locuteurs.

En fin de compte, cette étude des verbes modaux nous rappelle que chaque mot et chaque structure grammaticale possède une richesse de significations potentielles, façonnée par l'usage et le co(n)texte. En comprenant mieux ces mécanismes, nous pouvons apprécier davantage la complexité mais surtout la richesse du langage humain.

9. Bibliographie

- Abeillé, A., Godard, D., Delaveau, A., & Gautier, A. (2021a). *La Grande grammaire du français*. Arles : Actes Sud.
- Abeillé, A., & Koenig, J.-P. (2021a). La valence du verbe. In A. Abeillé, D. Godard, A. Delaveau, & A. Gautier (Eds.), *La Grande Grammaire du Français* (Vol. 1, pp. 148-180). Arles : Actes Sud.
- Abeillé, A., & Koenig, J.-P. (2021b). Les classes sémantiques des verbes. In A. Abeillé, D. Godard, A. Delaveau, & A. Gautier (Eds.), *La Grande Grammaire du Français* (Vol. 1, pp. 180-208). Arles : Actes Sud.
- Abeillé, A., Koenig, J.-P., Godard, D., & Bonami, O. (2021b). Qu'est-ce qu'un verbe ? In A. Abeillé, D. Godard, A. Delaveau, & A. Gautier (Eds.), *La Grande Grammaire du Français* (Vol. 1, pp. 127-147). Arles : Actes Sud.
- Achat, P., & Marty, E. (2016). Analyse diachronique d'une « trajectoire » médiatique : le cas de la crise de la dette souveraine en Grèce de 2009 à 2015. In *JADT 2016 : 13ème Journées internationales d'Analyse statistique des Données Textuelle*, Rome.
- Anokhina-Abeberry, O. (2000). *Étude sémantique du nom abstrait en français*. Thèse de doctorat. Université de Paris 8 Vincennes-Saint-Denis : Paris.
- Anscombre, J.-C. (1995). Morphologie et représentation événementielle : le cas des noms de sentiment et d'attitude. *Langue française*, 105, 40-54.
<http://www.jstor.org/stable/41558708>
- Anscombre, J.-C. (2015). Verbes d'activité de parole, verbes de parole et verbes de dire : des catégories linguistiques ? *Langue française*, 186 (2), 103-122.
- ATILF. (1998-2023). *Base textuelle Frantext*. ATILF-CNRS & Université de Lorraine.
<https://www.frantext.fr/>
- Aubry, L., & Pamuksaç, A. (2024). La presse sportive comme terrain d'étude pour les inégalités liées au genre : étude textométrique sur corpus. In *JADT 2024 - 17es Journées internationales d'Analyse statistique des Données Textuelles*, Bruxelles.
- Austin, J. L. (1970). *Quand Dire, C'est Faire*. Paris : Editions du Seuil.
- Avanzi, M., Béguelin, M.-J., & Diémoz, F. (2016). De l'archive de parole au corpus de référence : la base de données orales du français de Suisse romande (OFROM). *Corpus* (15).
- Bally, C. (1965). *Linguistique générale et linguistique française*. 4e éd., revue et corrigée ed., Berne : Francke.

- Balvet, A., Barque, L., Condette, M. H., Haas, P., Huyghe, R., Marin, R., & Merlo, A. (2011). La ressource Nomage. Confronter les attentes théoriques aux observations du comportement linguistique des nominalisations en corpus. *Revue TAL*, 52 (3), 129-152. <https://halshs.archives-ouvertes.fr/halshs-01077819> (Ressources Linguistiques Libres)
- Barbet, C. (2012). Devoir et pouvoir, des marqueurs modaux ou évidentiels ? *Langue française*, 1, 49-63.
- Barbet, C. (2013). *Sémantique et pragmatique des verbes modaux du français : Données synchroniques, diachroniques et expérimentales*. Thèse de doctorat. Université de Neuchâtel : Neuchâtel.
- Barbet, C. (2015). L'évolution sémantique des verbes modaux : hypothèses à partir des emplois de devoir et pouvoir en français moderne et médiéval. *Journal of French Language Studies*, 25 (2), 213-237. <https://doi.org/10.1017/S0959269515000113>
- Baroni, M., & Ever, S. (2008). Statistical methods for corpus exploitation. In A. Lüdeling & M. Kytö (Eds.), *Corpus linguistics : an international handbook*. Berlin : De Gruyter.
- Barthélémy, J.-P., & Luong, X. (1998). Représenter les données textuelles par des arbres. *4e journées internationales d'analyse statistique des données textuelles, Université de Nice*, 49-71.
- Bassnett-McGuire, S. (1980). *Translation studies*. London : Methuen.
- Bassnett, S. (2011). *Reflections on translation*. Bristol : Multilingual Matters.
- Bendinelli, M. (2012). *Étude des auxiliaires modaux et des semi-modaux dans les débats présidentiels américains (1960-2008) : analyse qualitative et quantitative*. Thèse de doctorat. Université Nice-Sophia Antipolis : Nice.
- Benzécri, J.-P. (1973). *L'analyse des données*. Paris : Dunod.
- Bertels, A., & Speelman, D. (2012). La contribution des cooccurrences de deuxième ordre à l'analyse sémantique. *Corpus*, 11, 147-165. <https://doi.org/10.4000/corpus.2184>
- Biber, D. (2015). Corpus-Based and Corpus-Driven Analyses of Language Variation and Use In B. Heine & H. Narrog (Eds.), *The Oxford Handbook of Linguistic Analysis* (pp. 193-224). Oxford : Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199677078.001.0001>
- Biemann, C., Heyer, G., Quasthoff, U., & Richter, M. (2007). The Leipzig corpora collection-monolingual corpora of standard size. In *Proceedings of Corpus Linguistic*, Birmingham.
- Blanche-Benveniste, C. (2002). Auxiliaires et degrés de « verbalité ». *Syntaxe & Sémantique*, 3 (1), 75-97. <https://doi.org/10.3917/ss.003.0075>

- Blanche-Benveniste, C. (2003). Les formes grammaticales de réalisation des sujets et leur inégale représentation en français contemporain. In J.-M. Merle (Ed.), *Le sujet* (pp. 73-90). Gap : Ophrys.
- Blumenthal, P. (2008). Histoires de mots : affinités (s)électives. In *Congrès mondial de linguistique française*, Paris.
- Blumenthal, P. (2011). Odeur – évolution des profils combinatoires. *Langages*, 181 (1), 53-71.
<https://doi.org/10.3917/lang.181.0053>
- Blumenthal, P. (2012). Les implications de la volition : Types de procès, centrage du verbe et polyphonie. In M. Birkelund, G. Boysen, & P. S. Kierkegaard (Eds.), *Aspects de la Modalité* (Vol. 469, pp. 31-44). Tübingen : Max Niemeyer Verlag.
- Bogacki, K. (2017). Quelques remarques sur les lexèmes français contenant une composante volitive. *Roczniki Humanistyczne*, 65 (8), 47-58.
- Borillo, A. (1982). Deux aspects de la modalisation assertive : croire et savoir. *Langages*, 67, 33-53. <http://www.jstor.org/stable/41681927>
- Brown, P., & Levinson, S. C. (1987). *Politeness: Some universals in language usage*. Cambridge : Cambridge University Press.
- Brunet, E. (2011). *Hyperbase, Logiciel hypertexte pour le traitement documentaire et statistique des corpus textuels, Manuel de Référence*. Nice : Université de Nice.
- Brunet, E. (2012). Nouveau traitement des cooccurrences dans Hyperbase. *Corpus*, 11, 219-246.
- Busquets, J., & Denis, P. (2001). L'ellipse modale en français : le cas de devoir et pouvoir. *Cahiers de grammaire*, 26, 55-74.
- Büyükgüzel, S. (2011). Modalité et subjectivité : Regard et positionnement du locuteur. *Synergies Turquies*, 4, 139-151.
- Caron, J., & Caron-Pargue, J. (2003). A multidimensional analysis of French modal verbs pouvoir, devoir and falloir. In F. H. van Eemeren, C. A. Willard, & F. Snoeck Henkemans (Ed.) *Proceedings of the Fifth Conference of the International Society for the Study of Argumentation*, Amsterdam.
- Caudal, P. (2021). L'infinitif. In A. Abeillé, D. Godard, A. Delaveau, & A. Gautier (Eds.), *La Grande Grammaire du Français* (Vol. 1, pp. 1317-1335). Arles : Actes Sud.
- Chomsky, N. (2002). *Syntactic Structures*. Berlin, New York : De Gruyter Mouton.
<https://doi.org/doi:10.1515/9783110218329>
- Chu, X. (2008). *Les verbes modaux du français*. Paris : Ophrys.

- Colas-Blaise, M. (2022). Mode d'existence, modalité et modalisation : les apports de la sémiotique. *Signata. Annales des sémiotiques/Annals of Semiotics*, 13, 1-26.
- Condoravdi, C., & Lauer, S. (2016). Anankastic conditionals are just conditionals. *Semantics and Pragmatics*, 9, 1-69, Article 8. <https://doi.org/10.3765/sp.9.8>
- Culioli, A. (1990). *Pour une linguistique de l'énonciation : Opérations et représentations (Tome 1)*. Paris : Ophrys.
- Culioli, A. (1999a). *Pour une linguistique de l'énonciation : Domaine notionnel (Tome 3)*. Paris : Ophrys.
- Culioli, A. (1999b). *Pour une linguistique de l'énonciation : Formalisation et opérations de repérage (Tome 2)*. Paris : Ophrys.
- Dansereau, D. M. (2016). *Variations stylistiques : cours de grammaire avancée*. New Haven : Yale University Press.
- De Haan, F. (1997). *The interaction of modality and negation: A typological study (Outstanding dissertations in linguistics)*. New York : Garland Publishing, Inc.
- De Saussure, L. (2012). Modalité épistémique, évidentialité et dépendance contextuelle. *Langue française*, 1, 131-143.
- De Saussure, L. (2014a). Verbes modaux et enrichissement pragmatique. *Langages*, 193 (1), 113-126. <https://doi.org/10.3917/lang.193.0113>
- De Saussure, L. (2014b). Verbes modaux et sous-détermination. *Mélanges offerts à Jacques Moeschler: Online publication*.
- De Saussure, L. (2017). Why French Modal Verbs Are not Polysemous, and Other Considerations on Conceptual and Procedural Meanings. In J. Blochowiak, C. Grisot, S. Durrleman, & C. Laenzlinger (Eds.), *Formal Models in the Study of Language* (pp. 281-296). Cham : Springer.
- Deboës, E. (2002). *Analyse sémantique des verbes français : aperçu d'une représentation du sens par des structures logico-sémantiques*. Thèse de doctorat. Université Lyon 2 : Lyon.
- Dell'Oro, F. (accepté-a). Modality and Mood. In E. Aldridge, A. Breitbarth, K. E. Kiss, A. Ledgeway, J. Salmons, & A. Simonenko (Eds.), *The Wiley Blackwell Companion to Diachronic Linguistics (DiaCom)*. Hoboken : John Wiley & Sons.
- Dell'Oro, F. (accepté-b). Volition ascription to the addressee in a diachronic perspective: A unifying hypothesis for some post-volitional developments of Latin *uolo*. *Journal of Historical Pragmatics*, 26 (2), XXX-XXX.

- Dell'Oro, F. (soumis). A comparative approach to WANT-constructions as sources of reported-speech and reported-thought evidential markers: Latin uolo vis-à-vis its cognates German wollen and French vouloir. XXX, XX (XX), XXX-XXX.
- Dell'Oro, F. (2022). Verbs of motion and intermediate source domains of modality: the understudied case of Italian *occorrere* 'to be necessary, to be needed'. *Folia Linguistica*, 43 (1), 65-95. <https://doi.org/doi:10.1515/flin-2022-2017>
- Dendale, P. (1994). 'Devoir' épistémique, marqueur modal ou évidentiel ? *Langue française*, 102, 24-40. <http://www.jstor.org/stable/41559281>
- Dendale, P. (2000). Devoir épistémique à l'indicatif et au conditionnel : inférence ou prédiction ? In A. Englebert (Ed.) *Actes du 22e Congrès international de Linguistique et de Philologie Romanes*, Bruxelles.
- Desclés, J.-P. (1999). Au sujet de la catégorisation verbale. *Faits de langues*, 7 (14), 227-237.
- Desclés, J.-P. (2003). Interactions entre les valeurs de pouvoir, vouloir, devoir. *Aspects de la Modalité*, 469, 49-66.
- Diewald, G. (2002). A model for relevant types of contexts in grammaticalization In I. Wischer & G. Diewald (Eds.), *New reflections on grammaticalization* (pp. 103-120). Amsterdam & Philadelphia : John Benjamins.
- Dister, A., Naets, H., & Watrin, P. (2022). Les noms d'agent au féminin dans un corpus de presse. Etude comparée Québec-Belgique. In *JADT 2022: 16th International Conference on Statistical Analysis of Textual Data*, Naples.
- Diwersy, S., & Legallois, D. (2014). L'apport de la méthode des motifs aux analyses phraséologiques en discours. In *Approches théoriques et empiriques en phraséologie*, Nancy, France. <https://hal.science/hal-03546271>
- Diwersy, S., & Luxardo, G. (2016). Mettre en évidence le temps lexical dans un corpus de grandes dimensions : l'exemple des débats du Parlement européen. *JADT 2016 - proceedings of 13th international conference on statistical analysis of textual data*, Nice.
- Domínguez, F. N. (2004). Le verbe vouloir dans les locutions verbales (de la théorie guillaumienne à la théorie pragmatique). In *Isla abierta: estudios franceses en memoria de Alejandro Cioranescu: X Coloquio de la Asociación de Profesores de Francés de la Universidad Española*, La Laguna.
- Dostie, G. (2004). *Pragmaticalisation et marqueurs discursifs*. Bruxelles : De Boeck & Larcier : Duculot.

- Douay, C. (2003). Des modalités de l'interlocution au système des modaux. *Corela. Cognition, représentation, langage*, 1 (1), 1-23.
- Dubois, J. (1970). DICTIONNAIRE ET DISCOURS DIDACTIQUE. *Langages*, 19, 35-47.
<http://www.jstor.org/stable/41680760>
- Dubois, J., & Dubois-Charlier, F. (1997). *Les Verbes français*. Paris : Larousse.
- Dubois, J., Giacomo, M., Guespin, L., Marcellesi, C., Marcellesi, J.-B., & Mével, J.-P. (2012). *Le dictionnaire de linguistique et des sciences du langage*. Paris : Larousse.
- Ducrot, O. (1984). *Le Dire et le Dit*. Paris : Editions de Minuit.
- Ducrot, O. (1993). À quoi sert le concept de modalité? In N. Dittmar & A. Reich (Eds.), *Modalité et acquisition des langues* (pp. 111-129). Berlin : De Gruyter.
- El Cherif, W. (2011). *Vers une classification sémantique fine des noms d'agent en français*. Mémoire de master. Dalhousie University : Halifax, Nova Scotia.
- Evert, S. (2005). *The statistics of word cooccurrences*. Thèse de doctorat. Stuttgart University: Stuttgart.
- Evert, S. (2009). Corpora and collocations. In A. Lüdeling & M. Kytö (Eds.), *Corpus Linguistics. Volume 2* (pp. 1212-1248). Berlin : De Gruyter Mouton.
- Fantoli, M., & Vandersmissen, M. (2019). Partitionnements multiples de corpus: une lecture polyangulaire? L'exemple des bases latines du LASLA. *Corpus*, 19, 1-14.
- Fasciolo, M., & Lammert, M. (2014). Types de noms et critères définitoires. *Travaux de linguistique*, 69 (2), 7-10. <https://doi.org/10.3917/tl.069.0007>
- Firth, J. (1957). A synopsis of linguistic theory, 1930-1955. *Studies in linguistic analysis*, 10-32.
- Flaux, N., & Lagae, V. (2014). Syntaxe et sémantique des marqueurs modaux : présentation. *Langages*, 1, 3-15.
- Flaux, N., & Stosic, D. (2014). Les noms d'idéalités et la modalité : marquage d'une opposition. *Langages*, 1, 127-142.
- Flaux, N., & Van de Velde, D. (2000). *Les noms en français : esquisse de classement*. Paris : Editions Ophrys.
- Fløttum, K., Jonasson, K., & Norén, C. (2007). *On : pronom à facettes*. Bruxelles : De Boeck/Duculot.
- Fournier, N., & Fuchs, C. (1998). Place du sujet nominal et opérations de thématization. *Cahiers de praxématique*, 30, 55-88.
- Franckel, J.-J. (2017). Vouloir dire. *LINX*, 74, 39-66. <https://doi.org/10.4000/linx.1705>

- Franckel, J.-J. (2018). Je vois ce que tu veux dire. *Corela. Cognition, représentation, langage*, 16 (1), 1-23.
- François, J., Le Pesant, D., & Leeman, D. (2007). Présentation de la classification des Verbes Français de Jean Dubois et Françoise Dubois-Charlier. *Langue française*, 153 (1), 3-19. <https://doi.org/10.3917/lf.153.0003>
- Fujimura, I. (2005). La féminisation des noms de métiers et des titres dans la presse française (1988-2001). *Mots. Les langages du politique*, 78, 37-52.
- Gadet, F. (2021). La variation régionale des périphrases verbales. In A. Abeillé, D. Godard, A. Delaveau, & A. Gautier (Eds.), *La Grande Grammaire du Français* (Vol. 1, pp. 1273-1275). Arles : Actes Sud.
- Gambette, P., Gala, N., & Nasr, A. (2012). Longueur de branches et arbres de mots. *Corpus*, 11, 129-145.
- Gambette, P., Lechevrel, N., & Trotot, C. (2021). Valoriser des corpus littéraires numériques avec Wikisource : de la recherche à la pédagogie. In L. Barbe & M. Severo (Eds.), *Wikipédia, objet de médiation et de transmission des savoirs*. (Nouvelle édition [en ligne]). Nanterre : Presses universitaires de Paris Nanterre.
- Girard, G. (2003). La notion de sujet : une notion à définir. In J.-M. Merle (Ed.), *Le sujet* (pp. 29-40). Gap : Ophrys.
- Girard, G. (2004). La notion de sujet et la notion de complément. *Cercles*, 9, 38-52.
- Godard, D., & Jayez, J. (1996). Types nominaux et anaphores : le cas des objets et des événements. In *Anaphores temporelles et (in-)cohérence* (pp. 41-58). Leyde : Brill.
- Goldschmitt, S., & Landvogt, A. (2008). Vouloir citationnel. Chronique du changement métonymique d'un verbe modal français. In *Congrès Mondial de Linguistique Française*, Paris.
- Gómez-Jordana Ferary, S., & Anscombe, J.-C. (2015). Introduction : Dire et ses marqueurs. *Langue française*, 186 (2), 5-12. <https://doi.org/10.3917/lf.186.0005>
- Goossens, V. (2009). La polysémie des noms d'affect. *Revue Tranel (Travaux neuchâtelois de linguistique)*, 50, 147-161.
- Gosselin, L. (2010). *La validation des Représentations : Les modalités en Français*. Amsterdam : Rodopi.
- Grice, H. P. (1979). Logique et conversation. *Communications*, 30 (1), 57-72.
- Gries, S. T., & Berez, A. L. (2017). Linguistic annotation in/for corpus linguistics. In N. Ide & J. Pustejovsky (Eds.), *Handbook of linguistic annotation* (pp. 379-409). New York : Springer.

- Haas, P., Barque, L., Huyghe, R., & Tribout, D. (2022). Pour une classification sémantique des noms en français appuyée sur des tests linguistiques. *Journal of French Language Studies*, 1-30. <https://doi.org/10.1017/S0959269522000187>
- Haas, P., & Huyghe, R. (2010). Les propriétés aspectuelles des noms d'activités. In E. Moline & C. Vetters (Eds.), *Temps, aspect et modalité en français* (pp. 103-118). Leyde : Brill.
- Habert, B., Nazarenko, A., & Salem, A. (1997). *Les linguistiques de corpus*. Paris : Armand Colin.
- Harner, H. (2016). *Focus and the semantics of desire predicates and directive verbs*. Thèse de doctorat. Georgetown University : Washington, D.C.
- Harner, H., & Khemlani, S. (2020). A theory of bouletic reasoning. *CogSci 42nd Annual Conference*, Toronto.
- Heenen, F. (2016). Imparfait et modalité. *Milli Mála*, 8, 93-117.
- Heiden, S., Magué, J.-P., & Pincemin, B. (2010). TXM : Une plateforme logicielle open-source pour la textométrie - conception et développement. In *10th International Conference on the Statistical Analysis of Textual Data - JADT 2010*, Rome, Italy. <https://halshs.archives-ouvertes.fr/halshs-00549779>
- Heine, B. (2002). On the role of context in grammaticalization. In I. Wischer & G. Diewald (Eds.), *New reflections on grammaticalization* (pp. 83-101). Amsterdam & Philadelphia : John Benjamins.
- Herslund, M. (2003). Faillir et falloir : la création d'opérateurs modaux. *Aspects de la Modalité*, 469, 67-73.
- Huot, H. (1974). *Le verbe devoir: étude synchronique et diachronique*. Paris : Klincksieck.
- Hütsch, A. (2020). *L'usage des verbes modaux en français et en allemand : étude contrastive de la combinatoire adverbiale sous l'éclairage quantitatif*. Thèse de doctorat. Université de Neuchâtel : Neuchâtel.
- Huyghe, R. (2011). La dénotation spatiale des noms d'événements. *Lingvisticae investigationes*, 34 (1), 138-155.
- Huyghe, R. (2012). Noms d'objets et noms d'événements : quelles frontières linguistiques ? *Scolia*, 26, 81-103.
- Huyghe, R. (2013). Autonomie ou dépendance sémantique des noms d'événements en français. *Travaux de linguistique*, 2, 7-23.
- Huyghe, R. (2014). La sémantique des noms d'action : quelques repères. *Cahiers de lexicologie*, 105, 181-201.
- Huyghe, R. (2015). Les typologies nominales : présentation. *Langue française*, 1, 5-27.

- Huyghe, R., & Tribout, D. (2015). Noms d'agents et noms d'instruments : le cas des déverbaux en-eur. *Langue française*, 1, 99-112.
- Jacques, M.-P. (2016). Une linguistique outillée, pour quels objets ? *Histoire Épistémologie Langage*, 87-99. https://www.persee.fr/doc/hel_0750-8069_2016_num_38_2_3563
- Julian, B., & Thorne, J. P. (1969). The Semantics of Modal Verbs. *Journal of Linguistics*, 5 (1), 57-74. <http://www.jstor.org/stable/4175018>
- Kastberg-Sjöblom, M. (2002). *L'écriture de J. M. G. Le Clézio : une approche lexicométrique*. Thèse de doctorat. Université de Nice : Nice.
- Kerbrat-Orecchioni, C. (1980). *L'énonciation de la subjectivité dans le langage*. Paris : Armand Colin.
- Kissine, M. (2007). *Contexte et force illocutoire : vers une théorie cognitive des actes de langage*. Thèse de doctorat. Université libre de Bruxelles : Bruxelles.
- Koenig, J.-P., & Abeillé, A. (2021). L'omission des compléments du verbe. In A. Abeillé, D. Godard, A. Delaveau, & A. Gautier (Eds.), *La Grande Grammaire du Français* (Vol. 1, pp. 238-247). Arles : Actes Sud.
- Kratzer, A. (1981). The notional category of modality. *Words, worlds, and contexts*, 38, 74.
- Kronning, H. (1988). Modalité, politesse et concession : "Je dois dire que". In H. Nølke (Ed.) *Opérateurs syntaxiques et cohésion discursive. Actes du IVe Colloque International de Linguistique Slavo-Romane*, Copenhagen.
- Kronning, H. (1990). Modalité et diachronie : du déontique à l'épistémique. L'évolution sémantique de *debere/devoir*. *Onzième Congrès des Romanistes Scandinaves, Trondheim 13-17 août 1990*, Trondheim.
- Kronning, H. (1996). *Modalité, cognition et polysémie : sémantique du verbe modal devoir*. Acta Universitatis Upsaliensis : Uppsala.
- Kronning, H. (2001). Pour une tripartition des emplois du modal devoir. In P. Dendale & J. Van der Auwera (Eds.), *Les verbes modaux* (pp. 67-84). Leyde : Brill.
- Kuiper, K., Fromont, R., & Gerhard, D. (2018). Polysemy and word frequency: A replication. *Journal of Research Design and Statistics in Linguistics and Communication Science*, 4 (2), 144-155. <https://doi.org/10.1558/jrds.33751>
- Lafon, P. (1980). Sur la variabilité de la fréquence des formes dans un corpus. *Mots. Les langages du politique*, 1 (1), 127-165. https://www.persee.fr/doc/mots_0243-6450_1980_num_1_1_1008
- Larrea, P. (1997). Notions et opérations modales : pouvoir, devoir, vouloir. In C. Rivière & M.-L. Groussier (Eds.), *La notion* (pp. 156-165). Paris : Ophrys.

- Larreya, P. (2004). L'expression de la modalité en français et en anglais (domaine verbal). *Revue belge de philologie et d'histoire*, 82 (3), 733-762.
<https://doi.org/doi:10.3406/rbph.2004.4856>
- Larsen, R. (1984). *Quelques regards sur le pronom 'on'*. Mémoire de maîtrise. Oslo : Université d'Oslo.
- Le Querler, N. (1996). *Typologie des modalités*. Caen : Presses Universitaires de Caen.
- Le Querler, N. (2001). La place du verbe modal pouvoir dans une typologie des modalités. In *Les verbes modaux* (pp. 17-32). Leyde : Brill.
- Le Querler, N. (2004). Les modalités en français. *Revue belge de philologie et d'histoire*, 643-656. https://www.persee.fr/doc/rbph_0035-0818_2004_num_82_3_4850
- Lebart, L., & Salem, A. (1994). *Statistique textuelle*. Paris : Dunod.
- Leblanc, J.-M. (2015). Proposition de protocole pour l'analyse des données textuelles : pour une démarche expérimentale en lexicométrie. *Nouvelles perspectives en sciences sociales*, 11 (1), 25-63. <https://doi.org/https://doi.org/10.7202/1035932ar>
- Leblanc, J.-M. (2017). Approfondissement - En savoir plus sur l'analyse factorielle des correspondances. In É. Née (Ed.), *Méthodes et outils informatiques pour l'analyse des discours* (pp. 218-228). Rennes : Presses universitaires de Rennes.
- Leblanc, J.-M., Fleury, S., & Née, É. (2017). Quels outils logiciels et pour quoi faire ? In É. Née (Ed.), *Méthodes et outils informatiques pour l'analyse des discours* (pp. 123-161). Rennes : Presses universitaires de Rennes.
- Leeman-Bouix, D. (2002). *Grammaire du verbe français : des formes au sens : modes, aspects, temps, auxiliaires*. Paris : Nathan.
- Legallois, D. (2012). La colligation : autre nom de la collocation grammaticale ou autre logique de la relation mutuelle entre syntaxe et sémantique ? *Corpus*, 11, 31-54.
<https://doi.org/10.4000/corpus.2202>
- Lehmann, C. (2002). New reflections on grammaticalization and lexicalization. In I. Wischer & G. Diewald (Eds.), *New reflections on grammaticalization* (pp. 1-18). Amsterdam & Philadelphia : John Benjamins.
- Lehmann, C. (2015). *Thoughts on grammaticalization*. Berlin : Language Science Press.
- Léon, J. (2008). Aux sources de la « Corpus Linguistics » : Firth et la London School. *Langages*, 171 (3), 12-33.
- Lochard, G., & Boyer, H. (1998). *La communication médiatique*. Paris : Editions du Seuil.
- Malrieu, D., & Rastier, F. (2001). Genres et variations morphosyntaxiques. *Traitement Automatique des langues*, 42 (2), 548-577.

- Martin, R. (1988). Croire que p / penser que p. *Cahiers d'Études Hispaniques Médiévales*, 7, 547-554. https://www.persee.fr/doc/cehm_0180-9997_1988_sup_7_1_2152
- Martin, R. (2005). Définir la modalité. *Revue de linguistique romane*, 69 (273-274), 7-18.
- Martinez, W. (2008). Répulsions lexicales : expériences autour de la cooccurrence négative. *JADT 2008 - Proceedings of the 9th international conference on the statistical analysis of textual data*, Lyon.
- Mathieu-Colas, M. (2006). Les classes de verbes : syntaxe et sémantique. In B. e. S. M. Jacqueline, [Le traitement du lexique. Catégorisation et Actualisation]. *Le traitement du lexique. Catégorisation et Actualisation*, Sousse, Tunisia.
- Mayaffre, D. (2004). *Paroles de président. Jacques Chirac (1995-2003) et le discours présidentiel sous la Vème République*. Paris : Honoré Champion.
- Mayaffre, D. (2014). Plaidoyer en faveur de l'Analyse de Données co(n)textuelles. Parcours cooccurrence dans le discours présidentiel français (1958-2014). In *JADT 2014 : 12es Journées internationales d'Analyse statistique des Données Textuelles*, Paris.
- Mayaffre, D., & Luong, X. (2003). Les discours de Jacques Chirac (1995-2002). Représentation arborée et généalogie politique. *Histoire & mesure*, 18 (XVIII-3/4), 289-311.
- Mayaffre, D., Pincemin, B., & Poudat, C. (2019). Explorer, mesurer, contextualiser. Quelques apports de la textométrie à l'analyse de discours. *Langue française*, 3, 101-115.
- McEnery, T., & Hardie, A. (2013). The History of Corpus Linguistics. In K. Allan (Ed.), *The Oxford Handbook of the History of Linguistics* (pp. 727-745). Oxford : Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199585847.013.0034>
- Monte, M. (2011). Modalités et modalisation : peut-on sortir des embarras typologiques ? *Modèles linguistiques*, XXXII (64), 85-101. <https://doi.org/10.4000/ml.353>
- Montrichard, C. (2020). *La presse de tranchées : Un espace discursif de mise en scène d'un contre-discours combattant ?* Thèse de doctorat. Université Bourgogne Franche-Comté : Besançon.
- Montrichard, C., Bermúdez Sabel, H., Rossari, C., & Dell'Oro, F. (2022). Interroger la modalité en latin et en français : construction, annotation et exploitation de corpus. *La constitution de corpus en diachronie longue : Méthodologies, objectifs et exploitations linguistiques et stylistiques*, Grenoble.
- Narrog, H. (2012a). Modality and Subjectivity. In H. Narrog (Ed.), *Modality, Subjectivity, and Semantic Change: A Cross-Linguistic Perspective* (pp. 5-60). Oxford : Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199694372.003.0002>

- Narrog, H. (2012b). *Modality, Subjectivity, and Semantic Change : A Cross-Linguistic Perspective*. Oxford : Oxford University Press.
- Narrog, H. (2012c). *Modality, subjectivity, and semantic change: a cross-linguistic perspective*. Oxford : Oxford University Press.
- Narrog, H. (2016). The expression of non-epistemic modal categories. In J. Nuyts & J. Van der Auwera (Eds.), *The Oxford handbook of modality and mood* (pp. 90-116). Oxford : Oxford University Press.
- Narrog, H., & Heine, B. (2011). Introduction. In B. Heine & H. Narrog (Eds.), *The Oxford Handbook of Grammaticalization* (pp. 1-16). Oxford : Oxford University Press.
<https://doi.org/10.1093/oxfordhb/9780199586783.013.0001>
- Née, É., & Fleury, S. (2017). Constituer un corpus en trois scénarios. In É. Née (Ed.), *Méthodes et outils informatiques pour l'analyse des discours* (pp. 63-101). Rennes : Presses universitaires de Rennes.
- Née, É., Leblanc, J.-M., & Fleury, S. (2017). Chapitre IV : Compter dans les textes, quelles unités ? In É. Née (Ed.), *Méthodes et outils informatiques pour l'analyse des discours* (pp. 103-121). Rennes : Presses universitaires de Rennes.
- Née, E., Sitri, F., & Veniard, M. (2014). Pour une approche des routines discursives dans les écrits professionnels. *Congrès Mondial de Linguistique Française – CMLF 2014*, Berlin.
- Née, É., Sitri, F., & Veniard, M. (2016). Les routines, une catégorie pour l'analyse de discours: le cas des rapports éducatifs. *Lidil. Revue de linguistique et de didactique des langues*, 53, 71-93.
- Nølke, H. (1993). *Le regard du locuteur. Pour une linguistique des traces énonciatives*. Paris : Éditions Kimé.
- Nuyts, J. (2016). Analyses of the modal meanings. In J. Nuyts & J. Van der Auwera (Eds.), *The Oxford handbook of modality and mood* (pp. 31-49). Oxford : Oxford University Press.
- Nuyts, J., & Van Der Auwera, J. (2016). *The Oxford handbook of modality and mood*. Oxford: Oxford University Press.
- Oahey, D. (2022). What can a corpus tell us about lexis? In A. O’Keeffe & M. J. McCarthy (Eds.), *The Routledge handbook of corpus linguistics* (pp. 185-203). London and New York : Routledge.
- Pakpahan, B., Sembiring, P., Harianja, N., & Marni, A. (2019). The Meaning of Verb Falloir and other Impersonal Verbs in French and Its Equivalences in Indonesian Language.

- Pamuksaç, A., & Sanvido, L. (2023). Distinguer des genres textuels en français langue étrangère: étude textométrique de marqueurs subjectifs. *Synergies Espagne*, 16, 129-145.
- Pariollaud, F. (2008). *Verbes : questions de sémantique*. Rennes : Presses universitaires de Rennes.
- Patard, A. (2015). Les constructions modales comparatives faire mieux de, valoir mieux et falloir mieux. *Syntaxe et sémantique*, 16 (1), 123-156. <https://doi.org/10.3917/ss.016.0123>
- Perret, D. (1974). Les verbes « pouvoir » et « vouloir » dans les énoncés de proposition. *Langue française*, 21, 106-121. <http://www.jstor.org/stable/41558489>
- Pincemin, B. (1999). Construire et utiliser un corpus: le point de vue d'une sémantique textuelle interprétative. *Atelier Corpus et TAL: pour une réflexion méthodologique, Conférence TALN*,
- Pincemin, B. (2012). Hétérogénéité des corpus et textométrie. *Langages*, 187 (3), 13-26. <https://doi.org/10.3917/lang.187.0013>
- Portine, H. (2017). Du contexte à la situation au cotexte et à l'intertexte. *Essais. Revue interdisciplinaire d'Humanités*, 12, 17-32.
- Pottier, B. (1964). Vers une sémantique moderne. *Travaux de sémantique et de littérature*, 2, 107-137.
- Poudat, C., & Landragin, F. (2017). *Explorer un corpus textuel : Méthodes, pratiques, outils (Champs linguistiques. Recherches)*. Louvain-la-Neuve : De Boeck Supérieur.
- Pusch, C. D. (2007). Faut dire : variation et sens d'un marqueur parenthétique entre connectivité et (inter)subjectivité. *Langue française*, 154 (2), 29-44. <https://doi.org/10.3917/lf.154.0029>
- Recanati, F. (2004). *Literal meaning*. Cambridge : Cambridge University Press.
- Ricci, C. (2017). *Les emplois modaux du futur et de l'imparfait : analyse contrastive italien-français avec un regard sur la diachronie*. Thèse de doctorat. Université de Neuchâtel : Neuchâtel.
- Richard, I. (2005). Les marqueurs verbaux de l'obligation en français et en anglais dans les textes. *La langue, le discours et la culture en anglais du droit*, 27, 69-85.
- Riegel, M., Pellat, J.-C., & Rioul, R. (2008). *Grammaire méthodique du français*. Paris : Presses Universitaires de France.

- Rossari, C. (2012). Valeur évidentielle et/ou modale de faut croire, on dirait et paraît. *Langue française*, 1, 65-81.
- Rossari, C. (2020). L'exploitation du potentiel argumentatif des formes modales : analyse qualitative et quantitative. In G. Dostie & P. Larrivée (Eds.), *Représentations du sens linguistique. Modalité intra- et extra-phrastique* (pp. 15-38). Caen : Presses universitaires de Caen.
- Rossari, C., Chessex, J., Ricci, C., Walther, I., & Wandel, D. (2020). Distribution of modal expressions of possibility and necessity in three encyclopedias covering two diachronic spans (18th and 21st centuries). In *JADT 2020 : 15es Journées internationales d'Analyse statistique des Données Textuelles*, Toulouse.
- Rossari, C., Cojocariu, C., Ricci, C., & Spiridon, A. (2007). Devoir et l'évidentialité en français et en roumain. *Discours. Revue de linguistique, psycholinguistique et informatique. A journal of linguistics, psycholinguistics and computational linguistics*, 1, 1-17.
- Rossari, C., Montrichard, C., Pamuksaç, A., Ricci, C., & Sanvido, L. (2022a). Structures argumentatives et analyses statistiques : étude des combinatoires entre connecteurs et indications énonciatives. *Humanidades & Inovação*, 9 (4), 398-416.
- Rossari, C., Montrichard, C., & Ricci, C. (2022b). Pour une approche sémantique des connecteurs au-delà de leurs propriétés relationnelles : étude sur des variations génériques et diachroniques dans des corpus écrits. *SHS Web Conf.*, 138, 11016. <https://doi.org/10.1051/shsconf/202213811016>
- Rossari, C., & Ricci, C. (2021). L'emploi de bien, ben/bene dans les séquences concessives. Quelques hypothèses sur les différences entre français et italien à partir d'une démarche quantitative. *Studi Italiani di Linguistica Teorica e Applicata* (1/2021), 100-119.
- Rossari, C., Ricci, C., & Dolamic, L. (2018a). Le conditionnel appliqué à devoir/dovere et son potentiel argumentatif. *Langue française*, 4, 105-120.
- Rossari, C., Ricci, C., & Dolamic, L. (2018b). Les modalités de représentation du discours dans Diderot et d'Alembert en comparaison avec deux encyclopédies du 21ème siècle : Universalis et Wikipédia. *Revue Tranel (Travaux neuchâtelois de linguistique)*, 69, 75-98.
- Rossari, C., Ricci, C., & Siminiciuc, E. (2016). La représentation des sens modaux dans trois langues romanes : le français, l'italien et le roumain. Du qualitatif au quantitatif et retour. *Syntaxe et sémantique*, 17 (1), 93-113. <https://doi.org/10.3917/ss.017.0093>

- Salvadori, J., Barque, L., Haas, P., Huyghe, R., Lombard, A., Monney, M., Schwab, S., Tribout, D., & Wauquier, M. (2021). *The semantics of deverbal nouns in French: Annotation guide*. Fribourg : Université de Fribourg.
- Schmid, H. (1996). *Treetagger*.
- Schnedecker, C. (2018). Les noms généraux dénotant l'humain. Introduction et présentation. *LINX*, 76.
- Sctrick, R. (2024). *MODALITÉ, linguistique Encyclopædia Universalis*. Consulté le 23 juin 2024 sur <http://www.universalis-edu.com/encyclopedie/semantique/>
- Searle, J. R. (1972). *Les Actes de Langage : Essai de Philosophie du Langage*. Paris : Hermann.
- Searle, J. R. (1979). *Sens et expression : Etudes de théories des actes de langage*. Paris : Editions de Minuit.
- Siepmann, D., Bürgel, C., & Diwersy, S. (2016). Le Corpus de référence du français contemporain (CRFC), un corpus massif du français largement diversifié par genres. *SHS Web of Conferences*, Tours.
- Sitri, F. (2013). Une lecture événementielle du verbe "pouvoir" dans des rapports de travailleurs sociaux. In D. Londei, S. Moirand, A. Reboul-Touré, & L. Reggiani (Eds.), *Dire l'évènement. Langage, mémoire, société* (pp. 73-83). Paris : Presses Sorbonne Nouvelle.
- Sitri, F., & Barats, C. (2017a). Constituer un corpus en analyse du discours, un moment crucial. In É. Née (Ed.), *Méthodes et outils informatiques pour l'analyse des discours* (pp. 41-62). Rennes : Presses universitaires de Rennes.
- Sitri, F., & Barats, C. (2017b). Introduction : Pour comprendre ce qui va suivre. In É. Née (Ed.), *Méthodes et outils informatiques pour l'analyse des discours* (pp. 9-16). Rennes : Presses universitaires de Rennes.
- Stage, L. (2003). Les valeurs modales du futur et du présent. In M. Birkelund, G. Boysen, & P. S. Kjærsgaard (Eds.), *Aspects de la Modalité* (pp. 203-216). Tübingen : Max Niemeyer Verlag.
- Sueur, J.-P. (1977). À propos des restrictions de sélection : les infinitifs devoir et pouvoir. *Lingvisticae investigationes*, 1 (2), 375-409.
- Sueur, J.-P. (1979). Une analyse sémantique des verbes devoir et pouvoir. *Français (Le) Moderne Paris*, 47 (2), 97-120.
- Tasmowski, L., & Dendale, P. (1994). Pouvoir E : un marqueur d'évidentialité. *Langue française*, 102, 41-55. <http://www.jstor.org/stable/41559282>

- Touratier, C. (1996). *Le système verbal français : (description morphologique et morphématique)*. Paris : A. Colin.
- Tutin, A. (2013). Les collocations lexicales : une relation essentiellement binaire définie par la relation prédicat-argument. *Langages*, 189 (1), 47-63. <https://doi.org/10.3917/lang.189.0047>
- Valiukienė, V. (2016). *La multifonctionnalité et les effets de sens de nécessité des verbes modaux devoir et falloir et leurs équivalents en lituanien. L'étude basée sur le corpus français-lituanien/lituanien-français*. Thèse de doctorat. Vilnius : Vilniaus universitetas.
- Valiukienė, V. (2022). Du verbe plein à la composante pragmatique : le cas de falloir. *Verbum*, 13, 1-13. <https://doi.org/10.15388/Verb.29>
- Van de Velde, D. (2010). Le rôle « instrument » en question. *Travaux de linguistique*, 61 (2), 7-30. <https://doi.org/10.3917/tl.061.0007>
- Van der Auwera, J., & Plungian, V. A. (1998). Modality's semantic map. *Linguistic Typology*, 2 (1), 79-124.
- Varvara, R., Salvadori, J., & Huyghe, R. (2022). Assessing affix polyfunctionality through BERT-derived representations. *20th International Morphology Meeting (IMM20)*, Budapest.
- Varvara, R., Salvadori, J., & Huyghe, R. (sous presse). Lexical ambiguity in contextualized word embeddings. A case study of nominalizations. *Lingue E Linguaggio*, XXX (XXX), XXX.
- Vendler, Z. (1957). Verbs and times. *The philosophical review*, 66 (2), 143-160.
- Veniard, M., & Sitri, F. (2017). Problématiques d'analyse du discours et méthodes In É. Née (Ed.), *Méthodes et outils informatiques pour l'analyse des discours* (pp. 163-202). Rennes : Presses universitaires de Rennes.
- Vet, C. (1994). "Savoir" et "croire". *Langue française*, 102, 56-68. <http://www.jstor.org/stable/41559283>
- Vet, C., & de Swart, H. (2021). L'interprétation des temps verbaux. In A. Abeillé, D. Godard, A. Delaveau, & A. Gautier (Eds.), *La Grande Grammaire du Français* (Vol. 1, pp. 1239-1264). Arles : Actes Sud.
- Vet, C., De Swart, H., Veters, C., & de Mulder, W. (2021). Les expressions de temps, d'aspect et de mode. In A. Abeillé, D. Godard, A. Delaveau, & A. Gautier (Eds.), *La Grande Grammaire du Français* (Vol. 1, pp. 1227-1238). Arles : Actes Sud.

- Vetters, C. (2004). Les verbes modaux pouvoir et devoir en français. *Revue belge de philologie et d'histoire*, 82 (3), 657-671. <https://doi.org/doi:10.3406/rbph.2004.4851>
- Vetters, C. (2007). L'emploi «sporadique» de pouvoir est-il aléthique? In *Études sémantiques et pragmatiques sur le temps, l'aspect et la modalité* (pp. 63-78). Leyde : Brill.
- Vetters, C. (2012). Modalité et évidentialité dans pouvoir et devoir : typologie et discussions. *Langue française*, 1, 31-47.
- Vetters, C. (2021). Les verbes modaux. In A. Abeillé, D. Godard, A. Delaveau, & A. Gautier (Eds.), *La Grande Grammaire du Français* (Vol. 1, pp. 1349-1359). Arles : Actes Sud.
- Vetters, C., & Barbet, C. (2006). Les emplois temporels des verbes modaux en français : le cas de devoir. *Cahiers de praxématique*, 47, 191-214.
- Vivès, R. (1998). « Les mots pour le dire » : vers la constitution d'une classe de prédicats. *Langages*, 131, 64-76. https://www.persee.fr/doc/lgge_0458-726x_1998_num_32_131_2168
- Wandel, D. (2020). *Emplois propositionnels et énonciatifs des cadratifs à nom abstrait en fait, en réalité, en vérité, en pratique, et de leurs équivalents allemands*. Thèse de doctorat. Université de Neuchâtel : Neuchâtel.
- Williams, G. (2003). From meaning to words and back: Corpus linguistics and specialised lexicography. *ASp. la revue du GERAS*, 39-40, 91-106.
- Wittgenstein, L. (1961). *Tractatus logico-philosophicus : suivi de, Investigations philosophiques*. Paris : Gallimard.
- Yoon, J., & Gries, S. T. (2016). Corpus-based approaches to Construction Grammar: Introduction. In J. Yoon & S. T. Gries (Eds.), *Corpus-based approaches to Construction Grammar* (pp. 1-8). Amsterdam : John Benjamins.
- Zribi-Hertz, A., & Abeillé, A. (2021). Les constructions passives, neutres et impersonnelles. In A. Abeillé, D. Godard, A. Delaveau, & A. Gautier (Eds.), *La Grande Grammaire du Français* (Vol. 1, pp. 209-238). Arles : Actes Sud.
- Zufferey, S., & Moeschler, J. (2012). *Initiation à l'étude du sens. Sémantique et pragmatique*. Auxerre : Éditions Sciences Humaines. <https://doi.org/10.3917/sh.zuffe.2012.01>

