

Les verbes modaux du français : ressemblances et dissociations à l'aide de leurs associations avec les verbes *dicendi* et *putandi*

Aylin Pamuksaç^{1,*}

¹Chaire de Linguistique Française, Institut des Sciences du Langage, Université de Neuchâtel, 2000 Neuchâtel, Suisse

Résumé. Cet article a pour but de saisir les ressemblances et les divergences entre les quatre principaux verbes modaux du français (*devoir*, *pouvoir*, *falloir* et *vouloir*) dans une démarche quantitative. Les études qualitatives à leur sujet attestent une polyvalence d'utilisation au sein des modalités et pourtant, ces verbes diffèrent dans leur usage. Pour ce faire, nous comparons leurs associations avec des verbes de communication et d'activité mentale afin de relever des nuances d'usage lorsqu'ils modalisent des complémentations verbales qui elles aussi révèlent une attitude du locuteur. Ces associations sont illustrées à l'aide d'un indice de spécificité qui est le *Log-Likelihood* et une représentation graphique grâce à l'analyse factorielle des correspondances.

Abstract. French modal verbs: similarities and dissociations using their associations with the verbs *dicendi* and *putandi*. This article aims to capture the similarities and differences between the four main modal verbs of French (*devoir*, *pouvoir*, *falloir* and *vouloir*) in a quantitative approach. Qualitative studies on these verbs show a polyvalence of use within modalities, yet they differ in their usage. To this end, we compare their associations with verbs of communication and mental activity in order to identify nuances of use when they modalize verbal complementation which also reveal a speaker's attitude. These associations are illustrated with the help of a specificity score, the *Log-Likelihood*, and a graphical representation using factorial correspondence analysis.

1 Introduction

Les modalités font couler beaucoup d'encre dans les études linguistiques depuis au moins deux décennies Narrog (2012 : 1) et se concentrent souvent sur les verbes modaux. Cependant, malgré l'émergence des outils numériques permettant, par exemple, des recherches quantitatives sur des grands corpus, les études existantes sur la modalité sont pour la plupart qualitatives et traitent notamment les différentes valeurs d'un même verbe modal (cf. pour *devoir* et ses différentes valeurs modales : Dendale (1994, 2000) ; Kronning (1996) ; Rossari et al. (2007) ; Vettters et Barbet (2006) ; cf. pour les valeurs de *pouvoir* : Tasmowski

* Corresponding author : aylin.pamuksac@unine.ch

et Dendale (1994) ; cf. pour *vouloir* : Domínguez (2004) ; Franckel (2017) ; Landvogt et Goldschmitt (2008)). D'autres encore comparent deux verbes modaux qui peuvent avoir la même valeur modale (cf. pour les points communs et différences entre *devoir* et *pouvoir* : Barbet (2012) ; Sueur (1977, 1979) ; Vettters (2004, 2012) ; de même qu'entre *devoir* et *falloir* : Rivière (1984) ; Valiukienė (2016)).

Dans cet article, compte tenu de leur polysémie respective, nous allons étudier les quatre verbes modaux dans une seule et même analyse. En effet, un même verbe peut exprimer différentes modalités et une même modalité peut être exprimée par différents verbes modaux. Pour ce faire, nous allons nous concentrer sur leurs complémentations verbales, angle encore peu abordé pour différencier les verbes modaux entre eux¹.

Contrairement aux recherches citées ci-dessus, nous ferons appel à des méthodes quantitatives pour saisir les nuances et les ressemblances entre les verbes modaux et plus précisément en investiguant des corpus de grande taille. Les outils sollicités dans cette étude sont l'indice de *Log-likelihood* (désormais LL) et l'Analyse Factorielle des Correspondances (désormais AFC). Le LL nous permettra de distinguer les accompagnateurs spécifiques de chaque verbe, aussi appelés cooccurrents, et la seconde nous fournira un point de vue plus global quant à la proximité ou à l'éloignement de ces verbes.

Notre démarche est corpus *based* mais aussi *driven*ⁱⁱ : *based*, car notre hypothèse est que les verbes modaux ne se combinent pas de façon similaire selon leur complémentation verbale. Nous ne nous sommes concentrée que sur les accompagnateurs spécifiques de la catégorie grammaticale des verbes car les verbes modaux sont parmi les seuls, en plus des auxiliaires, à nécessiter un autre verbe afin de pouvoir être interprétés (Abeillé et al. 2021 : 131)ⁱⁱⁱ. Notre approche est aussi corpus *driven*, car les types de verbes compléments sur lesquels nous avons travaillé ont été sélectionnés grâce à nos recherches sur corpus (cf. §3.1). Nous nous arrêterons ainsi sur une série de verbes qui marquent une attitude de communication ou d'activité mentale. Ces catégories sémantiques ont été choisies en nous fondant sur un critère quantitatif, leur co-fréquence avec les verbes modaux. Cependant, notre sélection s'est aussi appuyée sur des arguments qualitatifs. Dans les faits, les verbes *dicendi* et *putandi* donnent aussi des indications sur la façon dont le locuteur se positionne par rapport à un certain contenu, tout comme les verbes modaux, qui, dans la conception large de la modalité, traitent « [...] tout le champ des attitudes adoptées par le locuteur vis-à-vis du contenu propositionnel de son énoncé. » (Gosselin 2010 : 6). Ainsi, nous comparerons comment les verbes modaux se combinent avec d'autres formes verbales qui modalisent l'énoncé.

Nous commencerons donc par démontrer que les typologies traditionnelles des modalités ne sont pas suffisantes pour saisir les nuances entre les verbes modaux. Nous continuerons notre article en étayant notre choix des cooccurrents verbaux que nous avons pris en compte. Nous présenterons ensuite nos résultats obtenus avec l'indice de LL et les AFC que nous aurons produites grâce aux divers outils dont nous disposons. Enfin, nous terminerons par une interprétation de ces résultats et ce qu'ils impliquent pour le nuancement des quatre verbes modaux du français.

2 Cadrage théorique

Lorsque nous lisons les diverses études produites sur les verbes modaux, nous pouvons constater qu'elles se fondent majoritairement sur des études qualitatives. Ces verbes sont classés dans diverses catégories modales et ces classifications peuvent comporter quelques variations selon les auteur·e·s. Ainsi, dans la tradition francophone, Gosselin (2010) propose par exemple six catégories modales alors que Le Querler (1996) s'arrête à cinq classes. Ces deux linguistes partagent quatre sortes de modalités :

- i. les **modalités déontiques** qui définissent « [l']obligatoire, l'interdit, le permis (ou non interdit) et le facultatif [...] » (Gosselin 2010 : 360) donc « ce qui doit être, conformément à une règle sociale, morale, etc. » (Le Querler 1996 : 41) ;
- ii. les **modalités épistémiques** qui indiquent « le domaine du certain, du douteux, du savoir, de la croyance. » (Le Querler 1996 : 41) ;
- iii. les **modalités axiologiques**, qui sont chez Gosselin des « [...] jugements de valeur de nature morale, idéologique et/ou légale [...] » (Gosselin 2010 : 343) caractérisées par des expressions comme *il est juste que, je désapprouve le fait que*, etc. et chez Le Querler simplement « [...] l'évaluation du contenu propositionnel. » (Le Querler 1996 : 42) ;
- iv. les **modalités bouliques** qui « [...] marquent la volonté du sujet. » (Le Querler 1996 : 42) et permettent ainsi d'exprimer des « [...] désirs [et] des souhaits. » (Gosselin 2010 : 351).

Le Querler (1996 : 54) ajoute les *modalités temporelles* avec la valeur sporadique de *pouvoir* qu'on peut trouver dans l'exemple suivant :

1. Les Alsaciens *peuvent* être obèses.

qu'on pourrait comprendre comme :

2. Les Alsaciens peuvent *parfois* être obèses.

Gosselin (2010 : 314) rapprocherait cette valeur des *modalités aléthiques* qui sont de l'ordre de la vérité objective, avec des jugements descriptifs. Ainsi, nous pourrions également y classer des énoncés comme :

3. Si tu lances une pièce en l'air, elle *doit* retomber.

La dernière catégorie de Gosselin concerne les modalités appréciatives. Il différencie celles-ci des modalités bouliques en précisant qu'elles « n'expriment pas directement les désirs eux-mêmes [...] : elles servent à dire le désirable, *i.e.* à évaluer les objets et les procès sous l'angle des désirs (ou des aversions) qu'ils sont susceptibles de susciter, ou, en termes classiques à « *mettre le prix aux choses* ». » (Gosselin 2010 : 332-333).

D'autres linguistes se limitent, comme les philosophes de l'antiquité, au *possible* et au *nécessaire*. C'est le cas de Vetter (2021), qui précise que ces deux notions peuvent être exprimées à l'aide de trois modalités qui sont les modalités *déontiques*, *épistémiques* ou encore *descriptives*. Ainsi, le verbe *devoir* peut désigner une obligation (4), une probabilité (5) ou une description (comme dans l'exemple 3 ci-dessus) :

4. Un enfant *doit* obéir à ses parents.
5. Il *doit* pleuvoir demain.

Nous pouvons relever qu'avec ce type de classification les verbes eux-mêmes n'ont pas de valeur modale propre mais peuvent prendre la couleur de plusieurs modalités. Ainsi, le verbe *falloir*, qui généralement est considéré comme déontique, peut aussi revêtir une valeur épistémique :

6. Il *faut* que Pierre soit bête pour croire cela.

En somme, si nous résumons la pensée de ces chercheurs, les catégories modales servent à définir les différents points de vue du locuteur sur l'énoncé. Comme nous l'avons remarqué avec les divers exemples qui ont illustré ces classifications, les verbes modaux du français

peuvent très bien se retrouver dans plusieurs de ces catégories. Par exemple, les verbes *devoir* et *falloir* ne sont pas essentiellement déontiques, et peuvent très bien exprimer une modalité épistémique (exemples 5 et 6). Le verbe *pouvoir* n'exprime pas qu'une possibilité mais aussi une modalité dynamique, aléthique ou descriptive selon les auteurs (exemple 1). Le verbe *vouloir*, quant à lui, moins étudié par les divers·e·s auteur·e·s, peut aussi refléter différentes modalités. Les plus évidentes sont les modalités bouliques ou appréciatives mais il peut très bien servir à donner un ordre et donc exprimer une modalité déontique :

7. Je *veux* que tu ranges ta chambre.

La valeur déontique démontre aisément notre propos. En effet, l'exemple 7 pourrait très bien être exprimé avec les autres verbes modaux en français :

8. Tu *dois* ranger ta chambre.
9. Il *faut* ranger ta chambre.
10. *Peux-tu* ranger ta chambre ?

Comme nous l'avons remarqué avec les divers exemples qui ont illustré ces classifications, ce ne sont pas les typologies classiques de la modalité qui arrivent à différencier et catégoriser les verbes modaux de façon distincte. Ainsi, indépendamment de leur polysémie, nous proposons de trouver un moyen de les singulariser ou de les rapprocher dans l'usage que les locutrices et locuteurs en font. C'est pourquoi nous avons fait appel à une approche textométrique afin de constater si certains verbes modaux sont préférés selon la complémentation verbale qui suit dans des énoncés attestés.

3 Méthodologie et premiers résultats

3.1 Les cooccurrents verbaux de *pouvoir*, *devoir*, *falloir* et *vouloir*

Comme susmentionné, nous avons décidé d'étudier les différences et les rapprochements entre les verbes modaux par rapport à leurs profils combinatoires qu'ils arborent avec d'autres verbes. Nous avons effectué une recherche préliminaire dans le logiciel TXM (Heiden et al. 2010)^{iv} avec un corpus composé d'articles de Wikipédia^v. Sur les 44 verbes cooccurrents significatifs^{vi} que *devoir*, *falloir*, *pouvoir* et *vouloir* ont en commun, ressortent les catégories sémantiques des verbes *dicendi* et *putandi* avec des occurrences comme *dire*, *parler* ou encore *croire* et *comprendre*. D'autres verbes de dire ou de pensée apparaissent seulement pour certains verbes modaux comme *demander* qui est spécifique uniquement à *devoir* et *pouvoir*, *songer* à *falloir* et *pouvoir*, et ainsi de suite.

3.2 Choix des verbes pour l'étude

Suite aux premiers résultats quantitatifs exposés en §3.1, nous avons décidé d'étudier les combinaisons des quatre verbes modaux avec des verbes de communication et des verbes psychologiques, manifestant une activité mentale^{vii}.

Il est tout aussi pertinent de les rassembler sur le plan qualitatif car, comme l'écrit Rabatel (2003 : 67), ces verbes sont interchangeables et donc comparables. En effet, dans son analyse, il propose une catégorie réunissant ces deux types de verbes qu'il nomme « verbes de procès mental », puisqu'on pourrait très bien les substituer les uns aux autres dans des énoncés comme :

11. Jean *dit* que Pierre est en retard.

12. Jean *pense* que Pierre est en retard.

Un autre argument qualitatif qui justifie le choix de ces verbes en particulier est leur nature proche des verbes modaux. En effet, comme ceux-ci, ils modalisent l'énoncé (Gómez-Jordana Ferary & Anscombe 2015 : 7) et « [...] 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. » (Anscombe 2015 : 103). Ainsi, étudier l'association des verbes modaux avec d'autres types de verbes qui attestent d'une certaine présence du locuteur nous paraît adéquat.

C'est à partir du *Dictionnaire des Verbes Français* (désormais LVF) (Dubois & Dubois-Charlier 1997) que nous avons sélectionné les verbes de cette étude. Nous avons consulté la liste des verbes de communication et d'activité mentale et avons sélectionné les suivants :

- **10 verbes de communication** : *affirmer, dire, écrire, évoquer, exprimer, expliquer, parler, noter, mentionner* et *répondre*.
- **9 verbes psychologiques** : *considérer, croire, douter, imaginer, juger, penser, regretter, savoir* et *supposer*.

Le tri de la liste du LVF s'est effectué sur plusieurs critères. Nous avons d'abord exclu de la classe ceux qui étaient polysémiques comme *comprendre* qui peut avoir le sens de « être constitué de » comme dans :

13. L'offre *comprend* la nuit à l'hôtel et le petit-déjeuner.

Nous avons ensuite choisi des verbes qui apparaissent comme cooccurrents significatifs de tous les verbes modaux comme *croire, dire, parler, écrire, expliquer* et d'autres qui sont spécifiques seulement d'un, de deux, ou de trois modaux : *imaginer* pour *falloir* et *pouvoir*, *douter* seulement pour *pouvoir*, etc.

Au critère quantitatif, nous désirions aussi ajouter une certaine neutralité sémantique des verbes. Pour les verbes de communication, nous nous sommes appuyée sur trois aspects cités par Abeillé et Koenig (2021 : 199) : nous avons omis les verbes « [...] impliquant un changement d'attitude mentale [...] » (*ibid.*) comme *avertir* ou *prévenir* et nous avons aussi laissé de côté ceux « [...] inventés avec l'arrivée de nouveaux moyens de communication : *télégraphier*, puis *téléphoner*, puis *faxer*, puis *mailer* [et les] verbes [qui] indiquent la manière de communiquer, comme *bégayer, chuchoter, murmurer*. » (*ibid.*).

En ce qui concerne les verbes d'activité mentale, nous nous sommes à nouveau fondée sur le critère de neutralité sémantique susmentionné. Nous confirmons aussi le critère quantitatif en nous appuyant sur une liste de verbes nommée « thought activity » chez Rossari et al. (2020 : 4) que les auteur·e·s ont déterminé sur la base de la co-fréquence de ces verbes avec *devoir*, non seulement dans un corpus, mais dans trois, tous de type encyclopédique. Les auteur·e·s ont également travaillé sur Wikipédia, mais avec des articles extraits en 2015, ainsi que sur Universalis, également de 2015. À cela s'ajoute une autre tranche synchronique avec l'Encyclopédie de Diderot et d'Alembert du 18^{ème} siècle. Certes, cette liste s'appuie seulement sur les cooccurrents de *devoir* mais des verbes comme *juger* et *considérer* sont adéquats pour cette étude et, surtout, conviennent aux critères pour les verbes de communication cités par Abeillé et Koenig (2021) ci-dessus.

4 Les co-fréquences entre verbes modaux et verbes de dire et de pensée

Pour étudier comment les verbes modaux se différencient ou se rapprochent entre eux, nous les avons observés dans deux corpus à visée informative : un corpus d'encyclopédie

d'écriture collective, Wikipédia 2019, et un corpus de presse nationale française, Le Monde 2010^{viii}. Travailler sur ces deux corpus nous permet de constater si les particularités des verbes sont propres aux types de texte ou stables dans les deux genres à visée informative étudiés.

Dans un premier temps, nous avons calculé leurs cooccurrences avec les verbes de communication et d'activité mentale sélectionnés. Rappelons que « [l]a cooccurrence est l'association statistiquement significative de deux items (en général deux mots) dans une fenêtre déterminée du texte [...] » (Mayaffre 2014 : 4). Pour notre part, la fenêtre est un empan de 5 mots, car 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. Nous avons effectué un travail manuel en triant chaque occurrence dans laquelle un verbe modal apparaissait avec un des verbes *dicendi* ou *putandi* sélectionnés. Nous avons ainsi pu omettre des faux résultats comme :

14. On **peut** retrouver ces maisons **dites** maisons « d'ouvrier » de par leur architecture singulière. (Wikipédia)

Dans l'exemple 14, le logiciel aurait pu compter ceci comme une cooccurrence alors que nous pouvons très bien constater que c'est le verbe *retrouver* qui est le complément verbal de *pouvoir*.

Pour calculer les attractions entre ces verbes, nous avons utilisé un indice de spécificité. Le principe de cet indice repose sur une dimension statistique qui « [...] est le plus souvent calculée par comparaison entre la fréquence attendue de la rencontre entre les deux items considérés, et la fréquence effectivement constatée dans le corpus. » (*ibid.*). Pour reformuler, l'indice de spécificité indique un écart par rapport à une probabilité de rencontre attendue entre deux termes et la rencontre effective de ces deux unités. Nous avons opté pour l'indice de LL^{ix}, car il permet également de mettre en exergue des spécificités négatives^x contrairement à l'indice de Lafon (1980), proposé notamment par le logiciel TXM.

Pour obtenir l'indice de LL des cooccurrences, nous avons comptabilisé le nombre de fois que les verbes de dire et de pensée apparaissent avec *devoir*, *falloir*, *pouvoir* et *vouloir* dans les deux corpus grâce au logiciel TXM. Une fois le nombre de ces cooccurrences brutes comptabilisé, nous avons utilisé le logiciel R^{xi} pour calculer le LL :

Tableau 1. Valeurs des indices de LL dans Wikipédia 2019.

Verbe complétif	LL de <i>devoir</i>	LL de <i>pouvoir</i>	LL de <i>vouloir</i>	LL de <i>falloir</i>
affirmer	-0,9012	14,6316	2,2638	-1,0561
dire	-2,1059	182,1795	750,8586	25,7909
écrire	-0,1352	165,5952	2,796	-0,5835
évoquer	-1,3832	14,2596	0,594	0,5375
exprimer	-0,3108	286,5647	2,2964	-1,0389
expliquer	0,0204	599,7527	1,6257	1,4965
parler	1,2369	116,2036	8,7774	19,9721
noter	-1,8084	441,8098	-5,3923	638,2938
mentionner	3,6998	38,2359	-2,3925	74,4304
répondre	87,2333	22,7611	-0,2357	9,8366
considérer	67,252	1082,2779	-3,757	68,9962
croire	0,7394	46,8308	13,885	34,6392
douter	-0,9286	20,9415	-0,252	7,2444
imaginer	-0,5374	98,1535	-1,0444	34,0065
juger	2,781	15,2973	-0,2074	-3,133
penser	0,4037	177,2089	0,4696	12,9968
regretter	-1,1607	1,1341	-0,315	1,9743
savoir	-1,0804	-0,0945	1,6301	127,7199
supposer	-0,3503	105,4014	-0,0147	25,5557

Tableau 2. Valeurs des indices de LL dans Le Monde 2010.

Verbe complétif	LL de <i>devoir</i>	LL de <i>pouvoir</i>	LL de <i>vouloir</i>	LL de <i>falloir</i>
affirmer	-8,7938	-5,8602	-8,1569	-12,9205
dire	-7,7616	628,1075	1276,6732	305,5927
écrire	-0,2361	20,1315	19,6971	-0,9723
évoquer	-1,889	0,1586	-0,1215	-1,011
expliquer	-8,5944	18,5182	-23,8028	0,8889
exprimer	17,4203	112,3354	14,3371	-1,4725
mentionner	12,9894	-0,0054	-2,2688	0,5356
noter	-16,8234	-8,8501	-13,0599	27,6333
parler	5,3007	147,7281	107,8124	81,5594
répondre	226,1551	64,6477	30,2731	11,5767
considérer	13,2773	173,1812	-1,9173	19,1629
croire	-10,0267	184,2561	919,5304	82,7681
douter	0,1455	45,5282	-0,5299	3,4281
imaginer	-10,7868	491,2879	-0,6189	40,8886
juger	20,6299	13,9767	3,0391	-5,9157
penser	1,8794	271,3061	-0,5273	46,8469
regretter	-0,2006	52,2993	-6,5367	-0,0239
savoir	-2,3444	-44,8731	118,5005	207,3193
supposer	-14,1173	3,57	-6,4191	-6,4441

Ces deux tableaux nous permettent de relever quelques différences et points communs notables entre les verbes modaux.

De façon surprenante, *devoir* et *falloir*, les verbes *a priori* les plus forts sur l'échelle de la nécessité, diffèrent considérablement l'un de l'autre par rapport à leurs cooccurents, et ce, quel que soit le corpus. *Falloir* attire toujours plus de verbes de communication et de pensée que *devoir*. *Falloir* possède d'ailleurs presque les mêmes verbes comme cooccurents

spécifiques dans les deux genres (*dire, noter, parler, etc.*) et montre ainsi une certaine régularité dans son fonctionnement. Ce n'est pas le cas de *devoir*, qui, dans Wikipédia, possède seulement deux cooccurents significatifs positifs (*répondre* et *considérer*) et cinq dans Le Monde 2010 (*exprimer, mentionner, répondre, considérer, juger*). Dans ce corpus, il présente même des attirances négatives avec les verbes *noter* et *supposer*. Si l'on observe de plus près les cooccurents significatifs de ces deux verbes, il nous semble que ceux avec *devoir* sont plus « actifs » que ceux avec *falloir*. En effet, *falloir* attire des formes comme *dire, parler, croire, imaginer, savoir* ou encore *supposer* alors que *devoir* attire des verbes comme *exprimer* ou *juger*.

Devoir est d'ailleurs le verbe modal qui attire le moins de verbes de notre sélection : dans Wikipédia, il en attire deux, tout comme *vouloir*, et dans Le Monde 2010 il en attire cinq.

Pouvoir est le verbe modal qui attire le plus de verbes *dicendi* et *putandi* et ce, quel que soit le corpus. Cependant, il est intéressant de noter quelques différences selon les corpus : les verbes de *dire* sont, dans une moindre mesure, moins attirés par *pouvoir* dans le corpus de presse qu'ils ne le sont dans le corpus encyclopédique. En effet, dans Wikipédia, tous les verbes de dire sont significativement attirés par *pouvoir* alors que la liste se limite à six verbes dans Le Monde 2010 : *dire, écrire, expliquer, exprimer, parler* et *répondre*. En cela, *pouvoir* se rapproche de *falloir* qui comporte aussi moins de cooccurents significatifs dans Le Monde 2010 que dans Wikipédia. Les deux autres verbes modaux ont montré une nette augmentation des affinités avec les verbes de communication et d'activité mentale dans le corpus de presse par rapport au corpus encyclopédique.

Pouvoir est aussi singulier sur un autre point : c'est le seul modal qui possède des co-féquences uniques avec certains verbes. En effet, *affirmer, évoquer, douter* et *regretter* sont spécifiques qu'avec *pouvoir* dans l'un ou l'autre des corpus ou dans les deux.

Nous avons résumé nos observations dans le Tableau 3 qui permet de comparer, en un coup d'œil, les cooccurrences significatives des verbes modaux avec les autres verbes sélectionnés pour cet article, et ce, dans les deux corpus :

Tableau 3. Tableau récapitulatif des attirances des verbes modaux avec les verbes *dicendi* et *putandi* dans les corpus Wikipédia (Wiki) et Le Monde 2010 (LM). Les cases avec un + représentent un résultat du LL au-dessus de 10,83 et les cases avec un -, en-dessous du seuil.

	Devoir Wiki	Devoir LM	Pouvoir Wiki	Pouvoir LM	Vouloir Wiki	Vouloir LM	Falloir Wiki	Falloir LM
affirmer	∅		+	∅	∅		∅	-
dire	∅		+	+	+		+	+
écrire	∅		+		∅	+	∅	
évoquer	∅		+	∅	∅		∅	
expliquer	∅		+		∅	-	∅	
exprimer	∅	+	+	+	∅	+	∅	
mentionner	∅	+	+	∅	∅		+	∅
noter	∅	-	+	∅	∅	-	+	+
parler	∅		+		∅	+	+	+
répondre	+		+		∅	+	∅	+
considérer	+		+		∅		+	+
croire	∅		+		+		+	+
douter	∅		+		∅		∅	
imaginer	∅		+		∅		+	+
juger	∅	+	+		∅		∅	
penser	∅		+		∅		+	+
regretter	∅		∅	+	∅		∅	
savoir	∅		∅	-	∅	+	+	+
supposer	∅	-	+	∅	∅		+	∅

En somme, ce tableau montre que, même si *devoir* et *falloir* semblent *a priori* proches sémantiquement parlant, ils diffèrent quant à leurs contextes d'utilisation. *Falloir* ressemble

ainsi plus à *pouvoir* dans sa façon de se comporter avec d'autres verbes qui attestent une certaine présence du locuteur. Nous pouvons également constater que *vouloir* manifeste une plus grande attraction avec des verbes de communication plutôt que d'activité mentale alors que pour *falloir*, c'est l'inverse. *Vouloir* et *devoir* sont d'ailleurs les verbes qui attirent le moins de verbes de communication et d'activité mentale mais ne partagent pas nécessairement les mêmes compléments : ainsi, tous deux désattirent *noter* dans Le Monde et montrent une spécificité positive dans ce même corpus avec *mentionner* et *exprimer*. *Vouloir* attire des verbes comme *dire* et *parler*, des verbes de communications très neutres dans leurs traits sémantiques, alors que *devoir* montre une attraction avec *mentionner* ou *juger* qui semblent montrer une plus grande implication du locuteur dans leur utilisation.

Le calcul de LL permet manifestement de faire émerger des similitudes et différences entre les verbes modaux qui n'auraient pas été possibles de voir avec les seules études qualitatives. Nous avons toutefois voulu étayer ces résultats en faisant appel à une autre méthodologie afin de dégager de manière plus visuelle les points communs ou les oppositions entre les verbes modaux.

5 Visualisation de la proximité ou de l'éloignement entre verbes modaux grâce aux AFC

L'AFC est une méthode statistique qui est très facile à saisir. Elle « [...] traite des tableaux de nombres et elle remplace un tableau difficile à lire par un tableau plus simple à lire qui soit une bonne approximation de celui-ci. » (Cibois 1983 : 5). En effet, grâce à cette méthode développée par Benzécri, expliquée dans son ouvrage de 1973, la visualisation des résultats devient plus limpide. Cette méthode consiste « [...] le plus souvent à croiser, dans un tableau rectangulaire (tableau de contingence) des textes d'un côté et des unités textuelles élémentaires que sont les mots (lemmatisés ou non) de l'autre. » (Mayaffre 2014 : 9) et donc à étudier la proximité et/ou l'éloignement entre plusieurs textes selon les mots qui les composent. Cette méthode repose sur un calcul mettant en lien des points-lignes et des points-colonnes : « *Les méthodes factorielles*, largement fondées sur l'algèbre linéaire, produisent des représentations graphiques sur lesquelles les proximités géométriques usuelles entre points-lignes et entre points-colonnes traduisent les associations statistiques entre lignes et entre colonnes. » (Lebart & Salem 1994 : 80).

Le principe reste identique pour la représentation des cooccurrences. Dans notre cas, les points lignes représentent les verbes *dicendi* et *putandi* et les points colonnes, les modaux. Cette méthode permet autant d'étudier l'éloignement que la proximité entre plusieurs unités : « Il s'agit, pour l'AFC, et dans le strict cadre des corpus textuels, 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).

Nous avons réalisé deux AFC^{xii}, une pour chaque corpus, mettant ainsi en exergue les cooccurrences brutes de chaque verbe modal avec chaque verbe de dire ou de pensée :

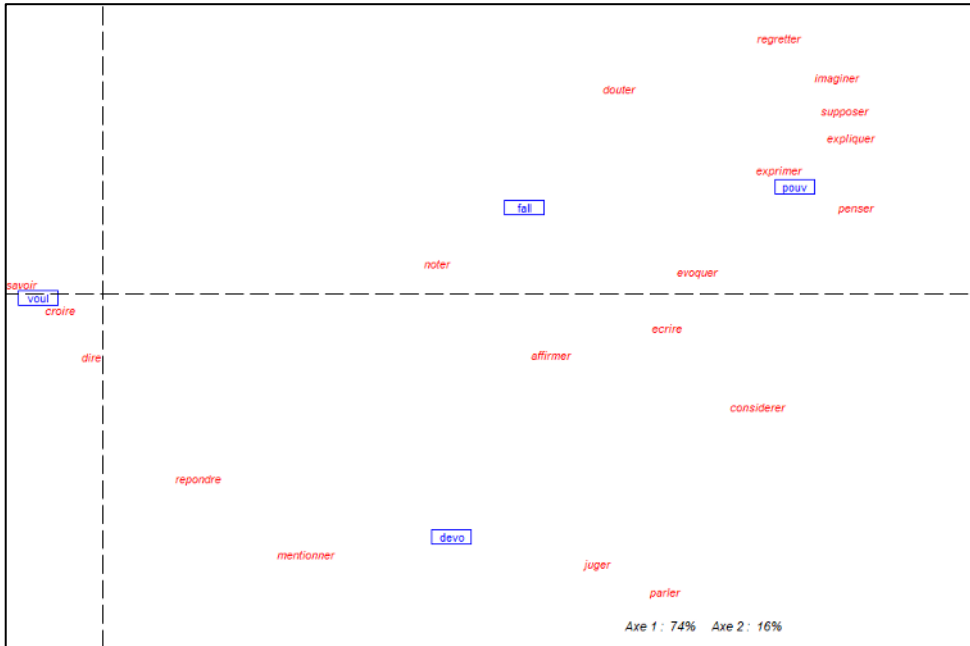


Fig. 1. AFC des cooccurrences entre verbes modaux et verbes *dicendi* et *putandi* dans Wikipédia 2019.

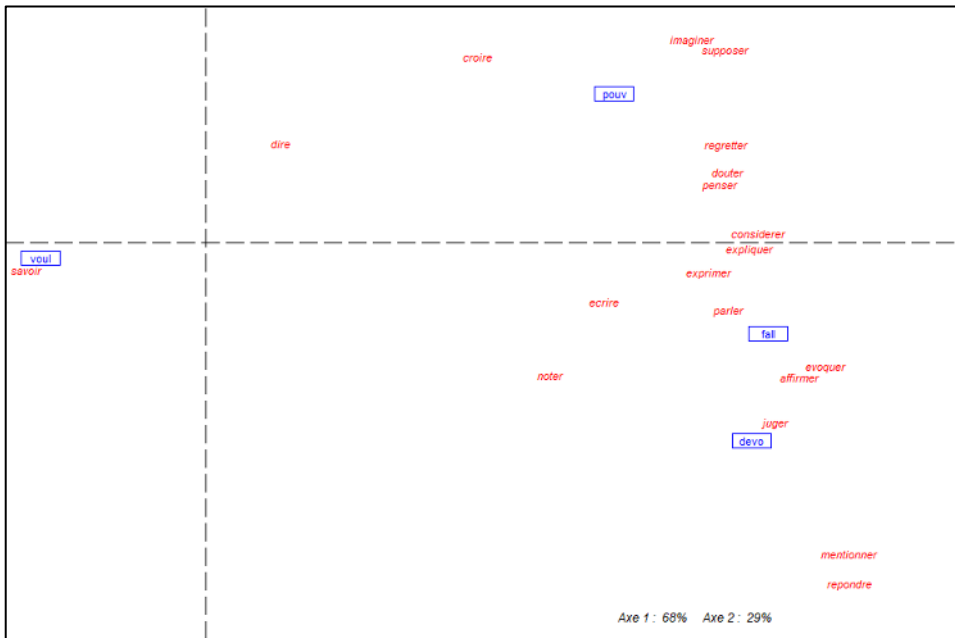


Fig. 2. AFC des cooccurrences entre verbes modaux et verbes *dicendi* et *putandi* dans Le Monde 2010.

Dans un premier temps, nous pouvons relever des points communs entre les deux AFC. En effet, peu importe le corpus, *vouloir* se trouve toujours à l'extrême gauche de l'axe 1 (horizontal) qui d'ailleurs représente la plus grande partie de l'information, quel que soit le corpus (74% pour Wikipédia et 68% pour Le Monde 2010). *Vouloir* se dissocie ainsi fortement de tous les autres verbes modaux.

Pour les observations liées à l'axe 2 (vertical), il faut montrer davantage de prudence. En effet, celui-ci représente bien moins d'informations que l'axe 1 (16% pour Wikipédia et 29% pour Le Monde). Relevons cependant que dans les deux corpus, *pouvoir* se situe en haut de l'axe et *devoir* à l'autre extrémité.

Il est intéressant de noter que *falloir* et *devoir* sont dans les deux corpus très proches sur l'axe 1 mais s'éloignent plus nettement sur l'axe 2. Sur ce dernier, *falloir* se rapproche de *pouvoir* et ceci est très marqué dans Wikipédia. *Falloir* se trouve d'ailleurs entre *devoir* et *pouvoir*, une fois sur l'axe 1 (Fig. 1) et une fois sur l'axe 2 (Fig. 2). Il semble ainsi y avoir une sorte de graduation qui est faite entre ces trois verbes, quel que soit le corpus.

En se penchant désormais sur la proximité des verbes modaux avec les autres verbes, nous pouvons aussi remarquer plusieurs régularités. Tout d'abord, *pouvoir* se rapproche beaucoup de verbes qu'on pourrait nommer « hypothétiques ». En effet, que ce soit dans l'encyclopédie ou dans la presse, il se trouve à proximité de verbes comme *imaginer*, *supposer*, *penser* ou encore *douter*. *A contrario*, *devoir* se rapproche de verbes « actifs » qu'on pourrait classer dans une macro-catégorie de verbes d'« actes d'énonciation » comme *juger*, *mentionner*, *répondre*, *affirmer*, *écrire*, etc.

Les observations que nous avons énumérées ici n'auraient pas pu être relevées à l'aide seule des indices de LL. En effet, grâce aux AFC, nous avons pu noter de manière plus claire une dichotomie entre *vouloir* et les autres verbes modaux ou encore une certaine graduation de *devoir* à *pouvoir* en passant par *falloir*. Cette graduation n'apparaît pas dans les études qualitatives citées en §2. Celles-ci admettaient que les modalités pouvaient être exprimées par différents verbes modaux et que ceux-ci étaient, par définition, polyvalents. Par contre, aucune progression de « force » ou de « degrés » ne semblaient émerger dans ces recherches-là.

5.1 Les similitudes entre *pouvoir* et *falloir*

Que ce soit avec l'indice de LL ou avec les AFC, *pouvoir* et *falloir* semblent montrer une certaine ressemblance. Dans la partie 4, nous avons relevé que *falloir* partageait plus de cooccurents spécifiques avec *pouvoir* qu'avec *devoir*. Pourtant, *devoir* et *falloir* sont les verbes déontologiques par excellence et pourraient ainsi partager des points communs autres que leurs traits sémantiques.

Dans les AFC, la proximité entre *pouvoir* et *falloir* était fortement marquée dans Wikipédia, car les deux verbes étaient avoisinants sur l'axe 1 qui comporte 74% des informations.

Afin de saisir ce rapprochement, nous avons suivi le conseil de Leblanc qui écrit que « [l']interprétation d'une analyse factorielle de correspondances ne peut se limiter à la description de la position des différents points sur le plan factoriel ; elle doit s'accompagner d'un retour au texte systématique. » (Leblanc 2015 : 53). Nous avons ainsi voulu voir en contexte comment le verbe *pouvoir* s'utilisait avec quatre verbes dont il est proche dans l'AFC de Wikipédia : *exprimer*, *expliquer*, *penser* et *supposer*. Dans un échantillon de 31 phrases prises au hasard, nous avons remarqué que *pouvoir* comporte sept fois le sujet *on* et trois fois *ce qui/ceci* qui le rapproche d'une utilisation impersonnelle du verbe près d'une fois sur trois.

Nous avons aussi observé cette régularité avec d'autres verbes avec lesquels il présentait un LL spécifique dans la partie 4 :

15. Parmi les prédécesseurs de Rousseau, **on peut noter** le Socrate de Platon.
16. Si les commanditaires sont importants dans la décision d'entreprendre la construction d'un ouvrage, **on ne peut imaginer** que cela ait été possible sans des maçons maîtrisant leur art.

Ce phénomène peut se rapprocher de « l'effet abstraitisant » de la combinaison *on peut* + *infinitif* relevé par Larsen (in Fløttum et al. 2007 : 22-23).

Nous avons effectué le même test avec *devoir*, pour voir si celui-ci revêtait de la même propriété. Nous avons observé les concordanciers lorsqu'il était suivi de *répondre*, *mentionner*, *juger* ou *parler*, ses verbes les plus proches dans l'AFC de Wikipédia (Fig. 1). Dans une sélection aléatoire de 31 phrases, *devoir* comporte six fois le pronom *on*, et présente ainsi une fois sur cinq une utilisation impersonnelle.

Nous avons aussi mené ces mêmes tests dans Le Monde 2010. *On* reste le pronom prédominant pour *pouvoir* mais le *nous* semble également important dans ce corpus. Ce *nous* apparaît souvent dans les citations et représente, certes, une collectivité, mais pas toujours une forme impersonnelle. Dans l'exemple (17), le *nous* représente les producteurs d'une télé-réalité qui parlent de deux candidats et dans (18) le *nous* semble plus impersonnel et représenter la totalité des personnes ou dans une moindre mesure, le lectorat de ce journal :

17. **Nous ne pouvions pas imaginer** qu'ils se disputeraient toute la journée !
18. C'est grâce aux reporters de guerre, aux risques qu'ils prennent, que **nous ne pouvons pas dire** que nous ne savions pas.

Dans ce même corpus, pour *devoir*, ni *on*, ni *nous* semble apparaître de façon significative avec ce verbe qui comporte des sujets humains, souvent repris par les pronoms *il* et *elle*.

Ainsi, cette proximité que nous avons pu observer dans les AFC et par les indices de LL pour *pouvoir* et *falloir* peut s'expliquer par une utilisation impersonnelle de *pouvoir*. La graduation de *devoir* à *pouvoir*, en passant par *falloir* s'expliquerait ainsi par une dichotomie entre un degré de nécessité fort et adressé avec *devoir*, en passant par une nécessité impersonnelle avec *falloir* pour terminer avec une possibilité, souvent impersonnelle, liée à *pouvoir*.

6 Conclusions

En somme, ce que notre étude quantitative démontre, c'est que, même si théoriquement les verbes modaux semblent pouvoir se substituer dans certains contextes à cause de leur polysémie attestée, ils sont, dans leur utilisation, divergents. *Devoir* et *falloir*, qui tous deux représentent une nécessité forte, ne se combinent pas avec les mêmes complémentations verbales. En effet, *falloir* se rapproche dans son utilisation de *pouvoir*, qui, pourtant, porte un sème de possibilité. Le peu d'études que nous pouvions trouver sur *falloir* ne pouvait pas prédire de tels résultats :

[...] French has basically two verbs to express strong and weak modality: *devoir* and *pouvoir*. Both are used in a deontic and an epistemic sense. In addition, French has the option of using the verb *falloir* and the impersonal construction *il faut que...* [...] to express strong deontic modality (but not epistemic modality). (de Haan 1997 : 95)

Cette proximité entre *falloir* et *pouvoir* peut s'expliquer par une utilisation impersonnelle de ce dernier, relevée dans les deux corpus, mais plus fortement dans l'encyclopédie Wikipédia.

Pris individuellement, chaque verbe possède des particularités : *falloir* semble stable selon le type de texte auquel il est associé, car ses verbes cooccurrents sont, pour la plupart, les mêmes, quel que soit le corpus étudié. *Vouloir* paraît aussi constant dans son utilisation avec les verbes *dicendi* et *putandi* dans les AFC, même si nous pouvons noter que le LL révèle que ce verbe s'associe davantage avec des verbes de *dire* dans la presse. *Devoir*, lui aussi, semble préférer la présence des verbes de communication et de pensée dans le corpus

Le Monde que dans Wikipédia. De plus, il est plus souvent associé à des verbes actifs comme *juger*, *affirmer*, *mentionner* ou *répondre*, que ce soit dans les AFC ou avec l'indice de LL.

Le fait de combiner deux méthodes statistiques et deux corpus de genres différents nous a ainsi permis de souligner de manière nette des régularités que ces verbes comportent. Avec l'indice de LL seul, nous n'aurions pas remarqué de manière si nette la proximité entre *pouvoir* et *falloir* et l'éloignement de *vouloir* par rapport aux autres verbes. Si nous avions étudié que Wikipédia, nous aurions pu penser que *devoir* et *vouloir* ne se combinaient que rarement avec les verbes *dicendi* et *putandi*, alors qu'ils affichent des cooccurrences significatives plus nettes dans Le Monde 2010.

En bref, grâce à cette étude, nous pourrions réenvisager une classification des verbes modaux en nous fondant sur un continuum allant d'une nécessité forte et adressée (*devoir*), qui passe par une forme forte mais impersonnelle (*falloir*) et qui se termine avec une possibilité, donc une nécessité faible et aussi impersonnelle (*pouvoir*). Ce continuum peut s'expliquer par les sujets grammaticaux avec lesquels *devoir* et *pouvoir* s'associent, mais aussi par une affiliation de *devoir* à des verbes sémantiquement plus marqués dans leur façon d'exprimer un dire et une proximité de *pouvoir* avec des verbes de pensée signalant une hypothèse. *Falloir* peut s'utiliser dans ces deux contextes, et ceci a été révélé tant par ses cooccurrences significatifs que par sa position dans les AFC. Quant à *vouloir*, il faudrait envisager une étude plus approfondie en le comparant à d'autres verbes qui peuvent exprimer une volonté ou un désir comme *souhaiter*, *désirer*, *espérer*, etc. pour obtenir un autre continuum, cette fois-ci centré sur les degrés de la volition.

Bibliographie

- Abeillé, A., & Koenig, J.-P. (2021). 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.
- 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.
- Barbet, C. (2012). Devoir et pouvoir, des marqueurs modaux ou évidentiels? *Langue française*, 1, 49-63.
- Benzécri, J.-P. (1973). *L'analyse des données*. Paris : Dunod.
- Cibois, P. (1983). *L'analyse factorielle*. Paris : Presses Universitaires de France.
- de Haan, F. (1997). *The interaction of modality and negation : A typological study (Outstanding dissertations in linguistics)*. New York : Garland Publishing, Inc.
- Dendale, P. (1994). "DEVOIR" ÉPISTÉMIQUE, MARQUEUR MODAL OU ÉVIDENTIEL? *Langue française*, 102, 24-40.
- Dendale, P. (2000). *Devoir épistémique à l'indicatif et au conditionnel: inférence ou prédiction?* Actes du 22e Congrès international de Linguistique et de Philologie Romanes, Bruxelles.
- Domínguez, F. N. (2004). *Le verbe vouloir dans les locutions verbales (de la théorie guillaumienne à la théorie pragmatique)*. Isla abierta: estudios franceses en memoria de Alejandro Cioranescu: X Coloquio de la Asociación de Profesores de Francés de la Universidad Española, La Laguna.
- Dubois, J., & Dubois-Charlier, F. (1997). *Les Verbes français*. Paris : Larousse.
- Fløttum, K., Jonasson, K., & Norén, C. (2007). *On : pronom à facettes*. Bruxelles : De Boeck/Duculot.
- Franckel, J.-J. (2017). Vouloir dire. *Linx*, 74, 39-66. <https://doi.org/10.4000/linx.1705>
- Gómez-Jordana Ferary, S., & Anscombre, J.-C. (2015). Introduction : Dire et ses marqueurs. *Langue française*, 186 (2), 5-12. <https://doi.org/10.3917/lf.186.0005>.
- Gosselin, L. (2010). *La validation des Représentations : Les modalités en Français*. Amsterdam : Rodopi.

- Kronning, H. (1996). *Modalité, cognition et polysémie : sémantique du verbe modal devoir*. Uppsala : Acta Universitatis Upsaliensis.
- 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.
- Landvogt, A., & Goldschmitt, S. (2008). 'Vouloir' citationnel. *Chronique du changement métonymique d'un verbe modal français*. Congrès Mondial de Linguistique Française, Paris.
- Le Querler, N. (1996). *Typologie des modalités*. Caen : Presses Universitaires de Caen.
- 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>.
- Mayaffre, D. (2014). *Plaidoyer en faveur de l'Analyse de Données co (n) Textuelles. Parcours cooccurentiels dans le discours présidentiel français (1958-2014)*. JADT 2014 : 12es Journées internationales d'Analyse statistique des Données Textuelles, Paris.
- Narrog, H. (2012). *Modality, Subjectivity, and Semantic Change : A Cross-Linguistic Perspective*. Oxford : Oxford University Press.
- Rabatel, A. (2003). Les verbes de perception en contexte d'effacement énonciatif : du point de vue représenté aux discours représentés. *Travaux de linguistique*, 1, 49-88.
- Rivière, C. (1984). Les équivalents anglais de «devoir» et «falloir». *Cahiers Charles V*, 6 (1), 7-27.
- 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)*. 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.
- Sueur, J.-P. (1977). À propos des restrictions de sélection : les infinitifs devoir et pouvoir. *Linguisticae 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*, 41-55.
- 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.
- Vetters, C. (2004). Les verbes modaux pouvoir et devoir en français. *Revue belge de philologie et d'histoire*, 82 (3), 657-671.
- 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.

ⁱ Sueur (1977) a étudié le contexte verbal dans lesquels s'insèrent *pouvoir* et *devoir* mais en tant que compléments infinitifs comme dans *il continue à devoir/pouvoir venir* (Sueur 1977 : 381) où les verbes modaux sont compléments de *continue* et non pas les verbes principaux de la phrase.

ⁱⁱ *driven* : « 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) et *based* : « La seconde approche est celle de la linguistique fondée sur le corpus [...], qui 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)

ⁱⁱⁱ Notons cependant que *vouloir* et *falloir* peuvent aussi se construire avec une complémentation nominale : *Je veux une glace* et *Il faut le passe sanitaire*. Les compléments nominaux n'ont pas été pris en compte ici car cette caractéristique n'est pas propre à tous les verbes modaux du français. Cependant, ils pourraient faire l'objet d'une étude similaire qui ferait appel à la même méthodologie que dans cet article.

^{iv} Logiciel libre de droit créé et développé à l'ENS de Lyon. Il est important de mentionner que dans le logiciel TXM, c'est l'indice de Lafon (1980) qui est utilisé. Cet indice est spécifique à partir de 2. Cela signifie que la cooccurrence observée a 1 chance sur 100 d'être due au hasard. Par souci de comparaison avec l'indice de LL que nous utilisons dans la suite de nos recherches, nous avons augmenté cette limite à 3 pour que nous puissions avoir des cooccurrences qui ont 1 chance sur 1000 d'être due au hasard.

^v Corpus composé de 18'715'455 mots tirés d'un échantillon de 10'570 articles choisis au hasard durant l'année 2019.

^{vi} Les verbes en commun dans l'ordre alphabétique sont : *aller*, ***appeler***, *appliquer*, *assurer*, *changer*, ***comprendre***, ***connaître***, *conserver*, *construire*, ***croire***, *détruire*, *développer*, ***dire***, *donner*, ***écrire***, *être*, *éviter*, ***expliquer***, *faire*, *garder*, *jouer*, *laisser*, *maintenir*, *mettre*, *obtenir*, *ouvrir*, ***parler***, *passer*, *payer*, *placer*, *prendre*, ***rappeler***, *reconnaître*, *rendre*, *rester*, *sauver*, *séparer*, *tourner*, *traiter*, *travailler*, *tuer*, *utiliser*, *vivre* et *voir*. (En gras, ceux que nous avons considéré comme faisant partie des catégories sémantique des verbes de communication ou d'activité mentale).

^{vii} Nous empruntons ces dénominations au dictionnaire électronique *Les Verbes Français* (LVF) élaboré par Dubois et Dubois-Charlier (1997) et disponible gratuitement en ligne depuis 2008 : http://rali.iro.umontreal.ca/LVF_DEM/LVF-DEM-req.html?dic=LVF.

^{viii} Corpus composé de tous les articles de l'année 2010 et comportant 17'895'009 mots.

^{ix} Le LL est un indice de cooccurrence. Il permet de constater si la cooccurrence entre deux formes est due au hasard ou non. Un indice est considéré comme significatif au-delà de 10,83. Cela signifie que la cooccurrence observée a 1 chance sur 1000 d'être due au hasard. En dessous de -10,83, nous pouvons parler de désattraction significative et entre ces deux seuils, d'indice de banalité.

^x Une spécificité positive indique que la cooccurrence est présente plus de fois qu'estimée selon la fréquence brute de chaque mot et de la taille du corpus. Une spécificité négative ne signifie pas que la cooccurrence n'existe pas dans le corpus, mais elle montre que la cooccurrence survient moins souvent que prévue par les fréquences des mots en question.

^{xi} R est un logiciel libre de droit développé par le projet GMU. Il permet d'utiliser le langage informatique pour le calcul statistique et les graphiques.

^{xii} Réalisées avec le logiciel Hyperbase, conçu et développé par Etienne Brunet et diffusé par le CNRS et l'Université Côte d'Azur.