

**LES VERBES D'INDICATION DU POINT
DE VUE DU LOCUTEUR
ETUDE CONTRASTIVE FRANÇAIS-
ITALIEN DANS UNE APPROCHE
DE LINGUISTIQUE DE CORPUS**

Thèse réalisée à la Faculté des Lettres et Sciences Humaines de l'Université
de Neuchâtel (Suisse), pour l'obtention du grade de docteur ès Lettres

par

LINDA SANVIDO

Membres du jury :

Corinne Rossari, professeure ordinaire et directrice de thèse, Université de Neuchâtel

Marion Carel, directrice d'études, EHESS, Paris

Anna-Maria De Cesare Greenwald, professeure, TU Dresden

Alain Kamber, professeur titulaire, Université de Neuchâtel

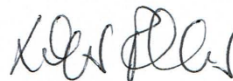
Thèse soutenue le 27 juin 2025



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 Carel, directrice d'études, EHESS, Paris ; Mme Anna-Maria De Cesare Greenwald, professeure, Technische Universität Dresden ; M. Alain Kamber, professeur titulaire, Université de Neuchâtel autorise l'impression de la thèse présentée par Mme Linda Sanvido en laissant à l'autrice la responsabilité des opinions énoncées.

Neuchâtel, le 27 juin 2025



Le doyen
Loris Petris

Remerciements

Ce travail est le fruit d'un long parcours que je n'aurais pas pu mener à bien sans le soutien de nombreuses personnes, que je souhaite remercier ici très chaleureusement.

Je tiens tout d'abord à exprimer ma profonde gratitude à ma directrice de thèse, Corinne Rossari. Merci, Corinne, pour les échanges très constructifs tout au long de ces années, pour ta confiance sans faille, ton enthousiasme et ta bienveillance.

Je remercie également les membres de mon jury de thèse, Marion Carel, Anna-Maria De Cesare et Alain Kamber. Je suis honorée que vous ayez accepté d'évaluer mon travail et vos remarques pertinentes ont grandement contribué à l'enrichir.

J'adresse ensuite toute ma gratitude à mes collègues de la chaire de linguistique française, et tout particulièrement à Claudia. Merci pour ces longues discussions où tu m'as appris tant de choses, non seulement en linguistique, mais aussi sur la vie en général. Tu n'as pas juste été une collègue formidable, tu es devenue une amie précieuse.

Merci aux collègues de PAM7 pour les collaborations, les repas partagés et les discussions dans les couloirs. Merci également aux professeur·es et collègues de l'ILCF : si je suis arrivée jusque-là aujourd'hui, c'est aussi grâce à vos enseignements et à vos encouragements.

Enfin, je voudrais remercier toutes celles et ceux qui ont été proches de moi au cours de ces cinq dernières années. Merci à mes ami·es, surtout à Ophélie, toujours présente pour célébrer les victoires et me reconforter dans les moments difficiles. Merci à ma famille, qui me soutient les yeux fermés depuis mes tout premiers pas en Suisse, et aux personnes qui sont venues l'agrandir, à Peseux et à Colombier. E grazie mille surtout à toi, Alexandre : tu m'as appris qu'il suffit de regarder dans les yeux de ceux qui t'aiment pour s'apprécier à sa juste valeur.

Résumé

Cette thèse s'intéresse à un paradigme de verbes d'indication du point de vue du locuteur, à savoir des formes qui ont pour caractéristique commune de présenter un état de fait comme relevant de la voix du locuteur. Le but est d'explorer leurs usages discursifs et leurs fonctions sémantiques et pragmatiques, dans une perspective contrastive entre français et italien et en adoptant une méthodologie mêlant analyses quantitatives et qualitatives. Les paradigmes des formes sont les suivants : *je crois que, je pense que, je trouve que, je considère que, je suis sûr-e que, je suis convaincu-e que, il me semble que, je dis que* pour le français et *credo che, penso che, ritengo che, sono sicuro/a che, sono convinto/a che, mi pare che* pour l'italien. L'approche adoptée permet de mettre en lumière des tendances d'usage de ces formes grâce à l'analyse statistique de leurs cooccurrents spécifiques et à l'établissement d'un profil combinatoire précis pour chaque verbe. Les analyses proposées permettent ainsi de dégager des régularités d'usage des verbes d'indication de point de vue, tout en révélant des distinctions intra- et interlinguistiques.

Mots-clés : verbes d'indication de point de vue, modalité, sémantique, linguistique de corpus, linguistique contrastive.

Abstract

This thesis focuses on a paradigm of verbs indicating the speaker's point of view, i.e. forms that share the characteristic of presenting a state of affairs as belonging to the voice of the speaker. The aim is to explore their discourse uses and their semantic and pragmatic functions, in a contrastive perspective between French and Italian and by adopting a methodology mixing quantitative and qualitative analyses. The paradigms of the forms are as follows: *je crois que, je pense que, je trouve que, je considère que, je suis sûr-e que, je suis convaincu-e que, il me semble que, je dis que* for French and *credo che, penso che, ritengo che, sono sicuro/a che, sono convinto/a che, mi pare che* for Italian. This approach highlights trends in the use of these forms by employing statistical analysis of their collocations and the establishment of a precise combinatorial profile for each verb. The proposed analyses identify regular patterns in the use of verbs expressing a point of view, while also uncovering distinctions both within and across languages.

Keywords: verbs indicating the speaker's point of view, modality, semantic, corpus linguistics, contrastive linguistics.

Table des matières

Liste des figures.....	17
Liste des tableaux	19
Liste des abréviations et symboles	21
1. Introduction	23
1.1 Objet d'étude.....	24
1.2 Les verbes « d'indication de point de vue » : pourquoi une nouvelle étiquette ?...26	
1.3 Intérêt de l'approche de linguistique de corpus.....	27
1.4 Articulation de la recherche.....	30
2. Perspectives théoriques sur l'utilisation des verbes d'indication de point de vue en discours	31
2.1 Quelques repères dans le domaine anglophone	32
2.1.1 Intérêt syntaxique des <i>parenthetical verbs</i>	33
2.1.2 <i>Les parenthetical verbs</i> comme marqueurs discursifs.....	35
2.2 Les verbes d'indication de point de vue dans la tradition française	37
2.2.1 La notion d'assertivité	38
2.2.2 La rection faible	41
2.2.3 Problématiques liées au contexte	45
2.2.4 Approches pragmatiques.....	47
2.2.5 Approches sémantiques	48
2.2.5.1 Je pense que, je crois que, je trouve que.....	49
Prédication originelle et prédication seconde	50
Développement ultérieurs.....	52
2.2.5.2 Je considère que.....	54
2.2.5.3 Et les autres verbes du paradigme ?.....	55
2.3 Les verbes d'indication de point de vue italiens.....	55
2.3.1 Venier (1991).....	56

2.3.2	Cignetti (2009)	56
2.3.3	Strik Lievers (2012).....	57
2.3.4	Lo Baido (2024)	58
2.3.5	Etudes syntaxiques	59
2.4	Approches comparatives	60
3.	Pour une analyse quantitative des verbes d'indication de point de vue : outils et méthode	65
3.1	Choix des corpus.....	66
3.2	Constitution des corpus et outils d'analyse	68
3.3	Sélection des paradigmes des verbes d'indication de point de vue	72
3.3.1	Extraction du paradigme des verbes français.....	76
3.3.2	Extraction du paradigme des verbes italiens.....	82
3.3.3	Optimisation du repérage des verbes dans nos corpus	88
3.3.4	Quelques réflexions sur les fréquences et la représentation des deux paradigmes dans les corpus.....	90
3.4	L'analyse cooccurentielle des verbes d'indication de point de vue	95
3.4.1	Fondements théoriques de l'analyse cooccurentielle.....	96
3.4.2	Choix méthodologiques pour l'établissement du profil combinatoire des verbes d'indication de point de vue	100
3.4.2.1	Etape préliminaire : définition des emfans et extraction des CS.....	101
3.4.2.2	Niveau macro : sélection des cooccurents spécifiques.....	102
	Réflexions théoriques pour la sélection des CS.....	102
	Création d'une macro-forme de CS	104
3.4.2.3	Niveau intermédiaire : classement des CS en catégories sémantiques.....	106
3.4.2.4	Niveau micro : analyse des CS.....	107
4.	Présentation des résultats de l'analyse cooccurentielle : les verbes d'indication de point de vue français.....	109
4.1	Le niveau macro.....	109

4.1.1	Interprétation des échelles de mise en scène énonciative.....	111
4.1.2	Synthèse du niveau macro	114
4.2	Le niveau intermédiaire.....	114
4.2.1	Interprétation des graphiques du niveau intermédiaire	117
4.2.2	Synthèse du niveau intermédiaire	118
4.3	Le niveau micro	118
4.3.1	Points communs entre les deux corpus français.....	121
4.3.2	Interprétation des tableaux du niveau micro.....	122
4.3.2.1	Je crois que et je pense que	123
	Les associations avec les CS de la classe <i>Construction argumentative</i>	123
	Les associations avec les CS de la classe <i>Prise en charge</i>	124
	Les associations avec les CS de la classe <i>Evaluation</i>	127
	Les associations avec les CS de la classe Modalité	128
4.3.2.2	Je trouve que	132
4.3.2.3	Je considère que et je dis que	134
4.3.2.4	Je suis sûr·e que et je suis convaincu·e que.....	136
4.3.2.5	Il me semble que.....	137
4.3.3	Synthèse du niveau micro	138
5.	Présentation des résultats de l'analyse cooccurrentielle : les verbes d'indication de point de vue italiens.....	139
5.1	Le niveau macro.....	139
5.1.1	Interprétation des échelles de mise en scène énonciative.....	141
5.1.2	Synthèse du niveau macro	143
5.2	Le niveau intermédiaire.....	144
5.2.1	Interprétation des graphiques du niveau intermédiaire	146
5.2.2	Synthèse du niveau intermédiaire	147
5.3	Le niveau micro	147

5.3.1	Points communs entre les deux corpus italiens.....	151
5.3.2	Interprétation des tableaux du niveau micro.....	153
5.3.2.1	Credo che et penso che	154
	Les associations avec les CS de la classe <i>Construction argumentative</i>	154
	Les associations avec les CS de la classe <i>Prise en charge</i>	156
	Les associations avec les CS de la classe <i>Evaluation</i>	159
	Les associations avec les CS de la classe <i>Modalité</i>	161
5.3.2.2	Sono convinto/a che et ritengo che.....	167
5.3.2.3	Mi pare che, sono sicuro/a che, dico che	171
	Mi pare che	171
	Sono sicuro/a che	173
	Dico che.....	174
5.3.3	Synthèse du niveau micro.....	177
6.	Mise en perspective des résultats français et italiens.....	179
6.1	Les niveaux macro	179
6.2	Les niveaux intermédiaires.....	180
6.3	Les niveaux micro.....	181
6.3.1	Je crois que/credo che et je pense que/penso che.....	183
6.3.2	Je considère que/ritengo che	184
6.3.3	Il me semble que/mi pare che	185
6.3.4	Je suis convaincu·e que/sono convinto/a che et je suis sûr·e que/sono sicuro/a che	186
6.3.5	Je dis que/dico che.....	188
6.4	Synthèse du chapitre comparatif.....	188
7.	Conclusion.....	191
7.1	Bref retour sur les objectifs	191
7.2	Apport de la thèse	193

7.2.1	Champs d'applications	196
7.3	Limites et prolongements	198
8.	Bibliographie	201
8.1	Références principales.....	201
8.2	Dictionnaires, encyclopédies et traducteurs en ligne	213
8.3	Corpus et bases de données	214
8.4	Logiciels et outils.....	215
9.	Annexes.....	217
9.1	Paradigmes des formes.....	217
9.2	Classes sémantiques	219
9.3	Listes exhaustives de tous les cooccurrents spécifiques	221
9.4	Script utilisé dans RStudio	245

Liste des figures

Figure 1 : illustration du phénomène de <i>slifting</i> (Ross 1973).....	35
Figure 2 : classification des verbes assertifs faibles et assertifs forts de Borillo (1982).....	39
Figure 3: classification de Ducrot (1975).....	51
Figure 4 : RPCs relevées par Schneider (2007 : 86).....	61
Figure 5 : exemple de recherche dans TXM.....	78
Figure 6 : exemple d’affichage des cooccurrents en word.....	83
Figure 7 : sur et sous-représentations des verbes d’indication de point de vue dans les corpus français.....	92
Figure 8 : sur et sous-représentations des verbes d’indication de point de vue dans les corpus italiens.....	93
Figure 9 : exemple de tableau de contingence (ici pour le corpus PRES-FR).....	105
Figure 10 : échelle de mise en scène énonciative des verbes d’indication de point de vue dans PRES-FR.....	110
Figure 11 : échelle de mise en scène énonciative des verbes d’indication de point de vue dans LM.....	110
Figure 12 : spécificités entre les verbes d’indication de point de vue et les classes sémantiques des CS dans PRES-FR.....	115
Figure 13 : spécificités entre les verbes d’indication de point de vue et les classes sémantiques des CS dans LM.....	116
Figure 14 : cooccurrents spécifiques retenus pour chaque verbe d’indication de point de vue dans PRES-FR.....	119
Figure 15 : cooccurrents spécifiques retenus pour chaque verbe d’indication de point de vue dans LM.....	120
Figure 16 : échelle de mise en scène énonciative des verbes d’indication de point de vue dans REP.....	140
Figure 17 : échelle de mise en scène énonciative des verbes d’indication de point de vue dans PRES-IT.....	140

Figure 18 : classes sémantiques spécifiques aux verbes d’indication de point de vue dans PRES-IT.....	145
Figure 19 : classes sémantiques spécifiques aux verbes d’indication de point de vue dans REP.	146
Figure 20 : cooccurrents spécifiques retenus pour chaque verbe d’indication de point de vue dans PRES-IT.....	148
Figure 21 : cooccurrents spécifiques retenus pour chaque verbe d’indication de point de vue dans REP.....	150
Figure 22 : exemple de traduction sur <i>DeepL</i>	196
Figure 23 : Recensement des structures servant à introduire l’opinion personnelle dans les <i>Référentiels pour le français</i> (Delahaie et al., 2022).....	197

Liste des tableaux

Tableau 1 : détails des corpus.	70
Tableau 2 : paradigme des verbes d'indication de point de vue français.	81
Tableau 3 : paradigme des verbes d'indication de point de vue italiens.	86
Tableau 4 : tableau récapitulatif des CS communs aux verbes d'indication de point de vue dans les corpus français.	122
Tableau 5 : tableau récapitulatif des CS communs aux verbes d'indication de point de vue dans les corpus italiens.	153
Tableau 6 : tableau comparatif des CS communs aux verbes d'indication de point de vue dans les deux langues.....	182
Tableau 7 : verbes d'indication de point de vue français exclus sur la base des critères de fréquence et de remplacement par « à mon avis ».	217
Tableau 8 : verbes d'indication de point de vue italiens exclus sur la base des critères de fréquence et de remplacement par « secondo me ».....	218
Tableau 9 : composition des classes sémantiques du niveau intermédiaire (français).....	219
Tableau 10 : composition des classes sémantiques du niveau intermédiaire (italien).	220

Liste des abréviations et symboles

Sigles des corpus

PRES-FR = corpus des discours/allocutions présidentiels français

LM = corpus de presse nationale, *Le Monde*

ER = corpus de presse régionale, *L'Est Républicain*

PRES-IT = corpus des discours/allocutions présidentiels italiens

REP = corpus de presse nationale, *La Repubblica*

AD = corpus de presse régionale, *l'Adige*

Symboles utilisés

* énoncé non acceptable d'un point de vue sémantique

? acceptabilité douteuse de l'énoncé

CS cooccurrent(s) spécifique(s)

∅ absence de marqueurs

P, Q propositions

p proposition complétive

ADJ adjectif

ADV adverbe

KON connecteur

NOM nom

PRO:PER pronom personnel

VER verbe

1. Introduction

Le présent travail¹ s'intéresse à un paradigme de formes verbales qui seront appelées *verbes d'indication du point de vue du locuteur*. Plus précisément, il s'agit de constructions verbales utilisées avec un morphème sujet de première personne du singulier et introduisant une proposition complétive. Les expressions soumises à l'étude concernent deux langues, le français et l'italien, et sont les suivantes :

- *je crois que, je pense que, je trouve que, je considère que, je suis sûr-e que, je suis convaincu-e que, il me semble que, je dis que* pour le français ;
- *credo che, penso che, ritengo che, sono sicuro/a che, sono convinto/a che, mi pare che* pour l'italien.

Ces expressions ont toutes comme caractéristique commune le fait d'introduire un contenu en l'étiquant comme relevant de la voix du locuteur². L'objectif de ce travail est d'analyser comment le locuteur mobilise ces expressions dans son discours, en cherchant donc à comprendre leur usage en contexte, tout en identifiant les similarités d'usage intralinguistiques (entre différents verbes au sein d'une même langue) et interlinguistiques (entre le français et l'italien). Pour mener cette étude, nous exploitons des corpus contenant des types de textes différents (des allocutions et discours présidentiels et des articles de presse). Plusieurs questions de recherche se posent :

- Quels sont les contextes d'apparition privilégiés des verbes d'indication de point de vue ?
- Est-ce que ces contextes d'apparition privilégiés mettent en lumière des similarités et des différences sémantiques entre ces formes ?
- Est-ce que les propriétés observées pour chaque forme sont stables selon le corpus mobilisé ?

¹ Cette thèse de doctorat s'inscrit dans le cadre du projet de recherche financé par le Fonds national suisse (FNS) intitulé *L'ancrage argumentatif des formes modales. Etude sur corpus avec un éclairage comparatif entre français et italien*, dirigé par la prof. Corinne Rossari (n° de projet : 192247, URL : <https://data.snf.ch/grants/grant/192247>).

² Dans cette thèse, nous adoptons un langage inclusif. Lorsque certains termes proviennent d'un cadre spécifique et désignent des notions théoriques, nous adoptons la dénomination de l'auteur·trice : cela s'applique notamment à la notion de « locuteur » issue de la théorie de la polyphonie de Ducrot (1984, 2001).

- Est-ce que les propriétés observées pour les verbes français se retrouvent pour les verbes morphologiquement ou sémantiquement équivalents en italien ?

L'analyse repose sur des outils statistiques qui permettent d'extraire des tendances récurrentes : la mise en perspective du sémantisme de ces expressions avec l'environnement discursifs dans lequel elles apparaissent de manière statistiquement significative permettra de cerner des propriétés qui ne peuvent pas être mises en lumière par une approche purement théorique.

1.1 Objet d'étude

Les verbes d'indication du point de vue du locuteur qui sont analysés dans cette étude apparaissent dans une proposition Q qui introduit à son tour une proposition complétive p . Par souci de praticité, lorsque nous parlons de « verbes d'indication de point de vue », nous nous référons aux séquences entières (pronom de première personne³ + syntagme verbal + conjonction) comme celles illustrées en gras dans les exemples suivants :

- (1) **Je pense que** tu as raison.
- (2) **Il me semble que** Pierre n'était pas là hier.
- (3) **Credo che** il maltempo abbia causato molti danni.
Je crois que le mauvais temps a causé beaucoup de dégâts.
- (4) **Sono sicura che** il convegno sarà un successo.
Je suis sûre que la conférence sera un succès.

Dans chacun de ces emplois, la proposition complétive p apporte le contenu informationnel principale de l'énoncé, alors que le verbe d'indication de point de vue ne fait qu'introduire ce contenu. Recanati (1984) qualifie ces séquences de transparentes, car elles n'apportent pas de contenu informationnel supplémentaire, mais indiquent uniquement comment l'énoncé global doit être compris. En d'autres termes, le but de ces verbes est celui d'étiqueter le contenu de p comme relevant de la **voix du locuteur**. Par locuteur, nous entendons l'instance de discours théorisée par Ducrot (1984), à savoir :

³ Non explicité dans les séquences en italien, à l'exception dans l'expression *mi pare che*.

un être qui, dans le sens même de l'énoncé, est présenté comme son responsable, c'est-à-dire comme quelqu'un à qui l'on doit imputer la responsabilité de cet énoncé. C'est à lui que réfèrent le pronom *je* et les autres marques de la première personne. (Ducrot 1984 : 193)

Le cadre théorique proposé par Ducrot (1984, 2001) constitue la base la plus pertinente et la plus efficace pour analyser le fonctionnement de ces formes dans le contexte de cette étude. En effet, notre intérêt se porte sur la sémantique sous-jacente à ces verbes, sur leur utilisation par le locuteur (en tant que responsable de l'énoncé) afin d'introduire des contenus. Nous ne prenons pas en compte l'aspect psychologique du « sujet parlant » (Ducrot 1984 : 199), c'est-à-dire l'individu empirique qui produit concrètement l'énoncé.

A l'intérieur même de cette notion de locuteur, Ducrot (1984 : 199) distingue le « locuteur en tant que tel » (abrégé « L ») et le « locuteur en tant qu'être du monde » (abrégé « λ »). L est le locuteur qui se montre dans son énonciation, comme dans l'énoncé (5), alors que λ qualifie l'énonciation, comme dans l'énoncé (6) :

(5) Hélas !

(6) Je suis triste.

L'interjection *hélas !* dans (5) met en scène L dans son engagement énonciatif, car elle « présente son énonciation comme déclenchée par le sentiment qu'elle exprime » (Ducrot 1984 : 200). En d'autres mots, L colore sa propre parole de tristesse. Mais dans (6), c'est λ qui intervient : il qualifie son énonciation en décrivant la tristesse, qui n'est pas montrée et qui reste extérieure à l'énonciation.

Les verbes d'indication de point de vue qui ont été sélectionnés pour l'analyse sont ceux qui sont utilisés de façon prototypique pour faire entendre la voix du locuteur (qu'il se manifeste en tant que L ou en tant que λ). En d'autres termes, les études statistiques menées sur les corpus à notre disposition ont permis de mettre en évidence une liste de verbes qui accomplissent, plus que d'autres, cette fonction en discours. Les réflexions théoriques et les étapes méthodologiques qui ont permis cette sélection seront présentées plus en détail dans la section 3.3.

L'objectif de cette thèse est de parvenir à une différenciation sémantique des verbes d'indication de point de vue : qu'est-ce qui motive le choix du locuteur d'utiliser un verbe plutôt qu'un autre ? Quels sont les effets discursifs que ce choix implique ? Des études

antérieures ont déjà proposé des pistes pour établir une distinction sémantique, notamment les travaux de Ducrot (1975), Gosselin (2015) et Tuchais (2012), qui seront détaillés dans la section 2.2.5. D'autres recherches, telles que celle de Borillo (1982) présentée dans la section 2.2.1, ont tenté de différencier ces verbes sur la base de leur assertivité. Enfin, la majorité de la littérature sur ces verbes s'intéresse à leur emploi en incise (notamment les études sur les verbes parenthétiques et les recteurs faibles présentés dans 2.1 et 2.2.2), emploi qui ne sera pas abordé dans le cadre de ce travail.

1.2 Les verbes « d'indication de point de vue » : pourquoi une nouvelle étiquette ?

La notion de point de vue, que nous employons pour nous référer aux séquences verbales qui font l'objet de la présente thèse, est reprise de Ducrot (1984). L'auteur l'introduit lorsqu'il explique les fondements de sa théorie de la polyphonie linguistique :

[...] le locuteur, responsable de l'énoncé, donne existence, au moyen de celui-ci, à des énonciateurs dont il organise les points de vue et les attitudes. Et sa position propre peut se manifester soit parce qu'il s'assimile à tel ou tel des énonciateurs, en le prenant pour représentant (l'énonciateur est alors actualisé), soit simplement parce qu'il a choisi de les faire apparaître et que leur apparition reste significative, même s'il ne s'assimile pas à eux. (Ducrot 1984 : 205)⁴

Cette notion se retrouve également dans Carel (2023), qui explique :

Par l'énonciation, le locuteur décrit son point de vue par rapport au contenu, il manifeste une position par rapport au contenu. (Carel 2023 : 178)⁵

En d'autres termes, les séquences verbales que nous appelons « verbes d'indication de point de vue » sont des marques énonciatives qui décrivent le point de vue, l'attitude, l'opinion, la position du locuteur vis-à-vis du contenu de l'énoncé.

⁴ Nous soulignons.

⁵ Nous soulignons.

L'adoption de cette étiquette présente l'avantage de pouvoir regrouper des formes qui n'ont pas encore été traitées ensemble jusqu'à présent :

- Les catégories sémantiques des grammaires traditionnelles séparent d'un côté les verbes d'opinion comme *je pense que, je crois que, penso che, credo che* et de l'autre des verbes de parole comme *je dis que, dico che*.
- Les études sémantiques en linguistique théorique tendent à classer, parmi ce vaste paradigme des verbes d'opinion, ceux qui expriment un jugement de nature évaluative et appréciative, comme *je trouve que, je considère que, ritengo che* et ceux qui introduisent un jugement de type épistémique, comme *il me semble que, mi pare che*, ou encore ceux qui sont compatibles avec les deux (*je pense que, je crois que, penso che, credo che*).
- Les études sur les verbes qui peuvent être utilisés en emploi parenthétique ne comprennent pas certains verbes de notre paradigme, comme *je considère que, je dis que, ritengo che*. En revanche, elles incluent des verbes conjugués à d'autres personnes (*tu vois, diciamo*), ainsi que des formes impersonnelles (*il paraît*). Ces deux cas ne nous intéressent pas, car nous nous focalisons uniquement sur les formes faisant intervenir la voix du locuteur. Par exemple, nous avons intégré *il me semble que* dans notre paradigme de verbes, mais nous avons exclu *il semble que* car dans le premier cas, le locuteur se porte responsable du contenu qu'il introduit, tandis que dans le second, sa voix n'est pas présente.

En conclusion, nous avons besoin de cette nouvelle classification pour nous concentrer précisément sur un paradigme de formes que nous estimons constituer un ensemble cohérent. En effet, il nous est impossible de nous référer à une classification existante, car nous nous intéressons à une classe de verbes que nous considérons comme formant un tout, malgré les divisions établies par les différentes approches théoriques détaillées dans le chapitre 2. Dans la section 3.3, nous démontrons que ces formes constituent un ensemble cohérent et qu'il existe un lien entre elles.

1.3 Intérêt de l'approche de linguistique de corpus

Les études qui se sont intéressées aux caractéristiques sémantiques propre à certains verbes d'indication de point de vue, comme celles qui ont différencié d'un côté *je pense que* et *je crois que* et de l'autre côté *je trouve que* (cf. section 2.2.5.1), ont le mérite d'avoir réussi à

mettre en évidence des propriétés spécifiques à ces expressions, en montrant qu'elles ne sont pas toujours substituables les unes aux autres. Cependant, ces analyses reposent souvent sur des exemples construits et décontextualisés, ce qui pose une limite importante à leur interprétation. En effet, le contexte joue un rôle essentiel dans la construction du sens d'une unité linguistique, comme le rappelle Vigier (2017) :

Les travaux de l'école « contextualiste » britannique issue de la *London school* [...], eux-mêmes influencés par les recherches philosophiques de Wittgenstein (1889-1951) sur le langage [...], ont largement contribué à développer et étayer l'idée suivant laquelle l'accès au sens d'une unité linguistique passe par l'exploration de son contexte d'apparition. On peut rappeler à ce propos la formule de J. Firth « *you shall know a word by the company it keeps* » (1957 : 11) qui n'est pas très éloignée de la phrase de L. Wittgenstein (*Investigations philosophiques*, 1953 : § 43) : « *La signification d'un mot est son usage dans le langage* ». (Vigier 2017 : 145-146)

Ducrot (1984) exprime également cette nécessité de prendre en compte les environnements d'apparition d'un mot ou d'un énoncé donné pour en saisir pleinement la signification :

La signification d'un énoncé, au regard de la linguistique, consisterait donc en une sorte de représentation condensée des associations dont il est susceptible dans l'usage (indiquant quels sont ses environnements, et quels autres énoncés ont les mêmes environnements que lui) [...]. La recherche sémantique vise donc à dire quels énoncés ont des contextes analogues dans un type déterminé de situation de parole et, de ce fait, semblent y avoir même fonction [...]. L'utilité, voire la nécessité, de cette démarche me semblent incontestables, surtout lorsqu'il s'agit d'étudier, à travers les relations de co-occurrence entre énoncés, les associations établies entre certains mots [...]. (Ducrot 1984 : 74-75)

Ces observations soulignent l'importance de considérer les environnements discursifs dans l'analyse linguistique des formes faisant l'objet de cette thèse. Bien que le recours aux corpus devienne de plus en plus courant pour l'analyse de ces verbes (citons par exemple les études de Andersen 2007, Blanche-Benveniste et Willems 2007, Schneider 2007, Rendulić et Kanaan-Caillol 2016, Angot 2021), ceux-ci ont principalement servi comme réservoirs d'exemples ou pour des études de fréquence, sans être véritablement exploités pour une analyse approfondie des entourages lexicaux à l'aide d'outils statistiques. Le manque d'études exploitant pleinement les possibilités offertes par la linguistique de corpus constitue l'une des principales motivations de cette recherche.

La linguistique de corpus se définit comme une « discipline empirique qui consiste à observer et analyser de manière quantitative des échantillons de langage rassemblé dans un format informatisé » (Zufferey, 2020 : 5). Les avancées informatiques des trente dernières années ont considérablement élargi l'accès à des corpus de textes, permettant ainsi un enrichissement des descriptions linguistiques. Ce type d'approche permet en effet de dépasser non seulement les limites d'une méthode strictement théorique (qui repose sur des exemples inventés et façonnés pour tester des hypothèses préconstruites), mais également celles des analyses exploitant les corpus en tant que simples réservoirs d'exemples, dont le dépouillement manuel est très laborieux et souvent peu représentatif sur le plan statistique.

L'utilisation des outils de la linguistique de corpus dans l'analyse des verbes d'indication de point de vue offre l'avantage de mettre en lumière leur fonctionnement à travers une analyse contextuelle appuyée sur des mesures statistiques. Cette approche permet notamment de mettre en perspective les différents types d'emplois de ces verbes afin d'identifier des tendances récurrentes, susceptibles de révéler à la fois des similitudes et des divergences dans leur usage. Plus précisément, cette étude vise à définir un « profil combinatoire » (Blumenthal 2002) pour chaque forme verbale, à savoir le portrait de la forme en question, établi sur la base des mots qui l'accompagnent de façon statistiquement significative, et que nous appellerons « cooccurrents spécifiques », abrégé « CS » (cf. section 3.4.1).

La construction du profil combinatoire repose sur une démarche en trois étapes, débouchant sur trois niveaux d'analyse : un niveau macro, qui aura pour but de classer les verbes d'indication de point de vue sur une échelle de force de la mise en scène énonciative ; un niveau intermédiaire, qui s'intéressera au type de cooccurrents spécifiques, en essayant de les classer d'un point de vue sémantique ; enfin, un niveau micro, qui prendra en compte les cooccurrents spécifiques un par un, pour étudier leur usage et les effets de leur interaction avec les verbes d'indication de point de vue. Au terme de ces étapes, un profil combinatoire sera établi pour chaque verbe, dans chaque corpus analysé. Les profils combinatoires ainsi construits serviront ensuite de base pour comparer les verbes entre eux, mettant en évidence les points communs ou les différences entre des formes qui sont sémantiquement très proches. Cette démarche sera appliquée de manière similaire aux verbes d'indication de point de vue italiens, pour ensuite procéder à une analyse comparative interlangue.

1.4 Articulation de la recherche

Ce travail de recherche s'articule en sept chapitres. Ce chapitre introductif est suivi par le chapitre 2, consacré à une présentation des études qui se sont intéressées aux verbes d'indication de point de vue faisant l'objet de cette thèse. Le chapitre 3 présente les outils et la méthodologie employés, en détaillant le choix et la constitution des corpus, les critères de sélection des verbes étudiés, ainsi que les étapes d'extraction et d'analyse de leurs profils combinatoires. Le chapitre 4 présente et analyse les résultats obtenus pour les formes françaises, selon les trois niveaux d'analyse (macro, intermédiaire, micro). Le chapitre 5 prend la même forme que le chapitre 4 en présentant et analysant à son tour les résultats pour les formes italiennes. Le chapitre 6 est consacré à une analyse comparative des résultats obtenus pour les deux langues, en mettant en lumière les correspondances et divergences les plus significatives. La conclusion au chapitre 7 rappelle les objectifs initiaux de la recherche, discute les apports majeurs de l'étude ainsi que ses limites, tout en ouvrant des perspectives pour de futurs travaux.

2. Perspectives théoriques sur l'utilisation des verbes d'indication de point de vue en discours

Les verbes objets de cette thèse ont suscité l'intérêt de nombreux chercheurs et chercheuses, qui les ont examinés dans des cadres théoriques très variés. Depuis le travail fondateur d'Urmson (1952), on les désigne généralement par le terme « parenthetical verbs » en raison de leur particularité à être syntaxiquement mobiles sans que leur sens change. En effet, des verbes comme *penser*, *croire* ou *trouver* présentent la proposition qu'ils accompagnent comme ayant une valeur de vérité affaiblie et ce indépendamment du fait qu'ils soient utilisés au début de l'énoncé (*Je pense/crois/trouve que Pierre a raison*), en incise interne (*Pierre a, je pense/crois/trouve, raison*) ou en position finale (*Pierre a raison, je pense/crois/trouve*). Dans le contexte francophone, c'est le terme « verbes recteurs faibles » qui s'impose depuis l'étude de Blanche-Benveniste (1989), mais les appellations de ce groupe de verbes sont très variées et les inventaires des formes parfois difficiles à circonscrire et pas toujours superposables. Rien que pour le français, Schneider (2009 : 88) cite une dizaine d'étiquettes :

À part *incise*, certainement le terme le plus répandu, on trouve *verbes d'opération* (Benveniste 1966 [1958]), *incidentes* (Dessaintes 1960 ; Mounin, 1974, p. 171 ; Riegel, Pellat & Rioul, 2002, p. 461), *modalisateurs d'assertion* (Borillo 1982), *verbes parenthétiques* (Cornulier 1978 ; Recanati 1984), *verbes recteurs faibles* (Blanche-Benveniste 1989 ; Blanche-Benveniste & Willems 2007), *propositions parenthétiques* (Andersen 1997), *proposition incidente à base de verbe épistémique* (Borillo, 2004, p. 33) et *clauses parenthétiques réduites* (Schneider 2007[...]).

L'état de l'art que nous présentons dans les pages suivantes n'a pas pour ambition d'être exhaustif : les travaux sur les verbes parenthétiques sont trop nombreux pour être tous abordés dans le cadre de cette étude. L'objectif est avant tout de mettre en évidence les propriétés communes généralement attribuées à cette classe plus ou moins homogène de verbes. Par la suite, nous nous concentrerons sur les recherches ayant spécifiquement exploré les propriétés sémantiques des verbes que nous avons sélectionnés, afin de mieux comprendre leurs spécificités individuelles.

Les pages qui suivent s'articulent de cette manière : la section 2.1 retrace les débuts de la réflexion sur les verbes parenthétiques anglais, en mettant en lumière les propriétés syntaxiques, sémantiques et pragmatiques qui ont conduit à les identifier comme une catégorie

distincte de verbes. La section 2.2 s'intéresse aux différentes théorisations sur cette classe de verbes dans le domaine francophone : nous présentons notamment la notion d'assertivité faible de Borillo (1982), la théorie sur les verbes recteurs faibles de Blanche-Benveniste (1989), ainsi que plusieurs études qui les abordent en tant que marqueurs discursifs (Apothéloz 2003, Andersen 2007, Schneider 2007). Ensuite, nous nous attarderons sur les études qui, sans regrouper systématiquement ces verbes au sein d'une même catégorie, ont mis en évidence les spécificités sémantiques qui distinguent certains d'entre eux. Enfin, la section 2.3 fait le point sur les recherches consacrées à ces verbes en italien.

2.1 Quelques repères dans le domaine anglophone

Nous devons à Urmson (1952) la découverte d'une classe de verbes aux propriétés communes appelés « parenthetical verbs ». Rentrent dans cette catégorie les verbes qui possèdent les trois propriétés suivantes : (i) ils sont utilisés à la première personne du singulier du présent de l'indicatif⁶ ; (ii) ils introduisent, au moyen de la conjonction *that*, une proposition complétive au présent de l'indicatif ; (iii) ils peuvent être utilisés en incise (au milieu ou à la fin de la proposition qu'ils accompagnent). Un exemple de verbe parenthétique est le verbe *suppose* :

(7) **I suppose that** your house is very old.

(8) Your house is, **I suppose**, very old.

(9) Your house is very old, **I suppose**.

(Urmson 1952 : 481)

Malgré cette mobilité syntaxique, le sens d'un verbe parenthétique ne change pas selon sa position dans l'énoncé : pour Urmson, les énoncés (7) à (9) sont sémantiquement équivalents, peu importe si *I suppose* intervient au début, au milieu ou à la fin. Le prédicat *suppose* a ainsi un emploi défini comme « pure parenthetical » (Urmson 1952 : 481), car dans les trois cas **il modifie et affaiblit la valeur de vérité de la complétive qu'il accompagne**. C'est donc le

⁶ Urmson (1952) s'intéresse uniquement aux verbes conjugués à la première personne du singulier de l'indicatif présent, mais les études ultérieures sur les parenthétiques prennent en compte également des verbes conjugués à d'autres personnes (notamment la deuxième personne du singulier et du pluriel), parfois à d'autres temps verbaux aussi (conditionnel, temps du passé).

contenu de la complétive qui est l'objet de l'assertion et les verbes parenthétiques ne font que donner des indications sur la façon dont l'assertion doit être comprise.

Urmson met en évidence trois groupes sémantiques de verbes parenthétiques : ceux qui donnent une orientation émotionnelle à l'énoncé (*regret, rejoice*), ceux qui fournissent des indications de pertinence logique (*deduce, presume, presuppose, assume*) et ceux qui indiquent le degré de fiabilité accordé à la proposition qu'ils introduisent (*know, believe, guess, suppose*). Le fonctionnement sémantique de ces verbes est ensuite mis en parallèle avec celui de certains adverbes, notamment évaluatifs pour le premier groupe (*happily, luckily, unfortunately*), évidentiels pour le deuxième (*admittedly, consequently, presumably*) et épistémiques pour le troisième (*certainly, possibly, probably*). En somme, pour être considéré comme parenthétique, un verbe doit respecter autant des critères syntaxiques (personne et temps verbal, construction avec une complétive, mobilité dans l'énoncé) et sémantico-pragmatiques (affaiblissement de la valeur de vérité de la complétive).

Bien qu'il ne les approfondisse pas, Urmson soulève des questions d'ambiguïté dans l'interprétation sémantico-pragmatique de ces prédicats qui seront aussi très débattues par la suite. Il évoque notamment certains contextes dans lesquels les verbes parenthétiques n'auraient pas un « pure parenthetical use » (Urmson 1952 : 481), car leur fonction discursive ne serait pas celle d'affaiblir la valeur de vérité de la complétive, mais plutôt de décrire un état psychologique du locuteur. Ainsi, un énoncé comme (7) pourrait tout à fait recevoir les deux interprétations selon le contexte discursif dans lequel il intervient. Ces questions de double interprétation des verbes parenthétiques anglais seront approfondies pour *I think* notamment (cf. Aijmer 1997, Simon-Vandenberghe 2000) et elles intéresseront également plusieurs chercheurs et chercheuses de la tradition française, depuis la différenciation problématisée par Benveniste (1966) entre emplois descriptifs et emplois mitigateurs (cf. section 2.2).

2.1.1 Intérêt syntaxique des *parenthetical verbs*

La particularité syntaxique de certains verbes à se déplacer de verbe principal à verbe en incise a été étudiée par un grand nombre d'études après Urmson. Dans le cadre de la grammaire générative, nous mentionnons notamment Jackendoff (1972), qui propose un parallèle très étroit entre verbes parenthétiques et **adverbes de phrase**, ce qui implique que les verbes parenthétiques ne sont pas considérés comme des verbes recteurs, mais comme un simple ajout à la construction. Verbes parenthétiques et adverbes partageraient alors plusieurs propriétés syntaxiques et sémantiques, notamment leur mobilité syntaxique et le fait qu'ils soient

speaker-oriented, car ils ne font que transmettre une attitude du locuteur par rapport à la proposition principale. Un exemple est représenté par le rapprochement entre le verbe parenthétique *I think* et des adverbes tels que *truthfully*, *honestly* et *sincerely* (Jackendoff 1972 : 99). Dans un cadre différent, celui de la grammaire fonctionnelle, Halliday (1985) rejoint les observations de Jackendoff (1972) en proposant un test syntaxique permettant de démontrer que le verbe parenthétique ne peut pas être un verbe recteur. En effet, pour lui, un énoncé comme *I think it is going to rain* est une simple variante de *It is probably going to rain* car le parenthétique *I think* est une expression de la modalité subjective du locuteur qui accompagne, au même titre que l’adverbe *probably*, la proposition principale *it is going to rain*. Afin de démontrer cela, Halliday (1985) utilise le test du *tag*, qui consiste à ajouter à l’énoncé la séquence *isn’t it ?* Ainsi, dans *I think it’s going to rain, isn’t it?* la reprise s’effectue bel et bien sur *it is going to rain* et non sur *I think*, ce qui confirmerait que les parenthétiques ne sont pas des verbes recteurs. Pour que *I think* soit considéré comme un énoncé régissant la complétive qu’il accompagne, la reprise doit se faire sur ce même énoncé au moyen de *don’t I ?* comme dans *I think it’s going to rain, don’t I ?* Halliday postule donc que même lorsqu’elle est antéposée, la proposition parenthétique ne régit pas forcément la proposition qui la suit, et démontre ainsi par des moyens syntaxiques ce que Benveniste décrivait pour le français (cf. la notion de « véritable énoncé », section 2.2).

Ces observations sont toutefois sujettes au débat. Ross (1973), par exemple, propose de traiter les parenthétiques comme des propositions principales, et ceci dans toutes leurs constructions possibles, que ce soient les emplois avec *that* (*I feel that Max is a Martian*) ou les emplois parenthétiques (*Max is a Martian, I feel*). Le rôle de proposition principale de *I feel* dans ces deux exemples est illustré à l’aide des arbres syntaxiques dans la Figure 1.

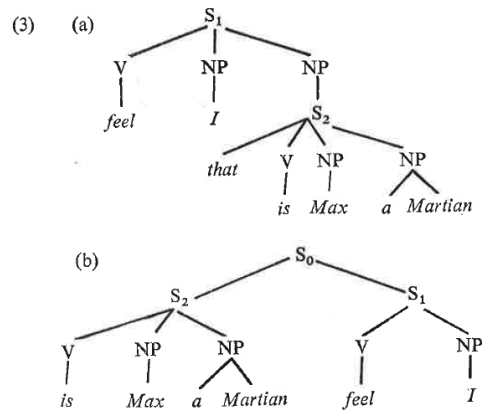


Figure 1 : illustration du phénomène de *slifting* (Ross 1973).

Ces variations syntaxiques sont un phénomène que Ross (1973) appelle *slifting* (pour *sentence lifting*) : cela se réfère précisément au déplacement du verbe, qui, dans les constructions parenthétiques, perd le *that* et vient suivre la complétive. Toutefois, dans les deux cas, les verbes appartiennent au même nœud S₁, qui régit S₂ dans le cas d'une construction avec *that* (3a) et qui reste indépendant dans le cas de l'emploi postposé (3b). Des arguments avancés à faveur du *slifting* sont le fait que seuls les verbes qui régissent des complétives avec *that* peuvent être utilisés de façon parenthétique et le fait que l'énoncé qui est accompagné par un parenthétique peut être pronominalisé en *so*. Ces réflexions amènent par conséquent cet auteur à considérer aussi les parenthétiques en incise comme des structures principales. Ces questions de portée syntaxique sont encore aujourd'hui assez débattues, mais pour ce qui concerne le français en particulier, les auteurs « semblent généralement peu enclins à admettre qu'un verbe ayant sémantiquement le rôle de mitiger la *que*-P puisse avoir au plan syntaxique la fonction de verbe principal recteur de la *que*-P » (Gachet 2015 : 231).

2.1.2 Les *parenthetical verbs* comme marqueurs discursifs

Dans une perspective pragmatique, les verbes parenthétiques ont été étudiés dans le cadre d'une plus vaste classe de marqueurs de discours appelés *hedges*. Cette notion a été popularisée par Lakoff (1973), qui s'intéresse à des expressions comme *sort of*, *kind of*, *more or less* qui sont employées pour exprimer de l'ambiguïté par rapport à un référent ou à une prédication. Ainsi, les *hedges* permettent au locuteur de se montrer prudent par rapport à son assertion. Bien qu'il ne parle pas explicitement de verbes parenthétiques, Lakoff remarque que certaines constructions comme *I suppose/guess/think that* peuvent être utilisées en tant que *hedges*. Par exemple, l'énoncé *I suppose that Harry is coming* est défini comme une *hedged*

assertion (Lakoff 1973 : 491), car le verbe *suppose* modifie la force du contenu propositionnel *Harry is coming*. Qu'on parle de modification de force du contenu propositionnel comme Lakoff (1973) ou de modification de la valeur de vérité de la complétive comme Urmson (1952), l'effet du parenthétique sur la proposition qu'il accompagne est le même : prendre de la distance. Le concept de *hedge* sera par la suite développé par Fraser (1975) – qui introduit le terme *hedged performative* pour décrire le comportement de certains verbes performatifs comme *promise, advise, admit* – et par Brown et Levinson (1990) – qui développent le concept de *speech act hedging* dans le cadre de leur théorie sur la politesse.

Certaines expressions parenthétiques, notamment *I think*, ont également été traitées dans le cadre des théories sur la grammaticalisation et la pragmaticalisation (cf. Traugott 1995, Aijmer 1997). Thompson et Mulac (1991) s'intéressent tout particulièrement aux constructions caractérisées par l'omission de la conjonction *that*, comme dans l'exemple suivant :

(10) I believe it's going to snow.

(Thompson & Mulac 1991 : 239)

L'originalité de cette étude est d'avoir démontré que l'alternance de construction (avec ou sans le *that*) n'est pas optionnelle en anglais, mais qu'il existe des conditions précises selon lesquelles le *that* peut ou ne peut pas être omis. Plusieurs contextes peuvent contraindre l'utilisation de la conjonction, par exemple lorsque le verbe parenthétique est modifié par un auxiliaire ou un adverbe. Inversement, les chances d'omission du *that* sont plus élevées quand le sujet du parenthétique est une première ou une deuxième personne du singulier de l'indicatif et que le prédicat relève de l'épistémique, en exprimant un degré plus ou moins haut d'adhésion du locuteur au contenu propositionnel de la complétive qu'il introduit. Dans le corpus de conversations pris en compte, les séquences qui ont le plus tendance à omettre le *that* sont celles construites avec les verbes *think* et *guess*. Ainsi, l'omission du *that* serait pour Thompson et Mulac (1991) une conséquence d'un processus de grammaticalisation de certains prédicats (comme *I think* ou *I guess*), qui deviennent presque des expressions figées. Ce phénomène est très systématique en anglais – pour Biber et al. (1999 : 681) l'omission du *that* relève de la norme –, mais cela n'est pas le cas pour le français, bien que l'alternance entre constructions avec ou sans *que* soit déjà attestée en ancien français (Glikman 2009) et continue de l'être dans certaines variétés de français moderne (cf. Gadet 1992, Martineau 1993, Wiesmath 2002, Akissi Boutin 2007, Avanzi 2009, Pustka 2010, Ferreira 2019).

2.2 Les verbes d'indication de point de vue dans la tradition française

Bien que l'ensemble des recherches sur les verbes parenthétiques anglais ait indéniablement influencé les réflexions sur les formes françaises correspondantes, c'est l'étude de Benveniste (1966) qui a posé, dans le domaine francophone, les premiers jalons d'une problématisation de cette question. Dans le cadre de sa théorie sur la subjectivité et en s'intéressant au sémantisme des verbes qui, utilisés à la première personne du singulier de l'indicatif présent, dénotent des « opérations mentales » (Benveniste 1966 : 264), Benveniste distingue entre prédicats ayant un emploi descriptif (comme *je sens que*) et prédicats ayant un emploi mitigateur (comme *je crois que*) :

En disant *je sens (que le temps va changer)*, je décris une impression qui m'affecte. Mais que se passera-t-il si, au lieu de *je sens (que le temps va changer)*, je dis : *je crois (que le temps va changer)* ? La symétrie formelle est complète entre *je sens* et *je crois*. L'est-elle pour le sens ? Puis-je considérer ce *je crois* comme une description de moi-même au même titre que *je sens* ? Est-ce que je me décris croyant quand je dis *je crois (que...)* ? Sûrement non. L'opération de pensée n'est nullement l'objet de l'énoncé ; *je crois (que...)* équivaut à une assertion mitigée. En disant *je crois (que...)*, je convertis en une énonciation subjective le fait asserté impersonnellement, à savoir *le temps va changer*, qui est la véritable proposition.

Ces réflexions transposent sur le plan sémantique le débat concernant la relation syntaxique de ces verbes avec la proposition qu'ils accompagnent : est-ce qu'ils la régissent toujours ? Pour Benveniste, ce n'est pas le cas. Recanati (1984) défendra également cette position en introduisant la notion de « transparence » :

Dans les exemples de Benveniste, [...] l'emploi non descriptif est caractérisé par la transparence de la proposition principale où figure le verbe à la première personne par rapport à la proposition enchâssée, qui est le « véritable énoncé » : la proposition principale n'ajoute rien au contenu de l'énoncé – à ce qui est dit par son moyen – elle indique seulement comment il doit être pris. Du fait de cette transparence de la proposition principale, l'énoncé global a à peu près la même valeur qu'aurait l'énoncé isolé de la complétive : dire **je crois que P**, **je suppose que P**, **je conclus que P** ou **je présume que P**, c'est, essentiellement, dire que P. (Recanati 1984 : 321)

Les verbes transparents permettent au locuteur de présenter le contenu *p* sous forme d'un « commentaire par lequel celui qui parle indique à son auditoire comment son énoncé doit être

pris » (Recanati 1984 : 324). Ainsi, les verbes utilisés avec l'emploi « non descriptif » (l'emploi mitigateur dont parle Benveniste) sont des espèces d'étiquettes relevant du métatexte, qui n'apportent pas d'information supplémentaire au contenu qu'ils introduisent : ils ont une fonction d'« indicateurs pragmatiques » (Recanati 1984 : 325). Cette transparence ne caractérise pas les verbes performatifs comme *remercier* ou *parier* : le fait de dire *je te remercie d'être venu* n'est pas une façon d'asserter *tu es venu*, de la même façon que dire *je parie six mille francs que Belle d'Azur arrivera en tête* n'est pas une façon de dire que Belle d'Azur gagnera (Recanati 1984 : 322). Contrairement aux verbes mitigateurs tels que *croire* ou *supposer*, les verbes performatifs ne sont pas employés pour présenter un contenu d'une certaine manière, mais participent eux-mêmes au contenu de l'énoncé. Recanati cite également la possibilité des deux interprétations des verbes transparents : des expressions comme *je crois que p* ou *j'affirme que p* peuvent avoir soit un emploi descriptif (description d'une croyance ou d'une affirmation), soit un emploi mitigateur (fonction d'indicateurs pragmatiques). Pour lui, cette ambiguïté est pragmatique et le fait de pencher pour l'une ou l'autre interprétation ne peut se décider qu'en contexte.

2.2.1 La notion d'assertivité

Les travaux de Borillo sur le concept de « modalité assertive » (Borillo 1976, 1982)⁷ permettent de montrer que, selon le verbe que le locuteur choisit d'utiliser pour présenter le contenu *p*, la valeur de *p* ne sera pas nécessairement la même. Borillo se réfère aux marqueurs tels que *je crois que* par le terme « modalisateurs d'assertion ». Elle les définit comme des verbes qui, « à la forme affirmative et renvoyant au locuteur, [...] attribuent ou conservent à la proposition qu'ils introduisent la valeur d'une assertion ou d'une quasi-assertion, même s'ils l'affaiblissent par la modalité épistémique dont ils sont l'expression » (Borillo 1982 : 33). Observons ces deux exemples, qui mettent en scène un contenu *p* (*il n'est pas coupable*) et le préfacent de deux manières différentes :

(11) J'affirme qu'il n'est pas coupable.

(12) J'ai l'impression qu'il n'est pas coupable.

⁷ Ses travaux sont largement influencés par la notion de *assertive predicate* de Hooper (1975), qui définit ainsi ces verbes : « assertive predicates imply that the speaker or subject of the sentence has an affirmative opinion regarding the truth value of the complement proposition » (Hooper 1975 : 95).

(Borillo 1982 : 33)

La proposition *il n'est pas coupable* dans (11) a une valeur d'assertion plus forte que la même proposition dans (12) : cela est dû au choix du verbe *affirmer*, qui véhicule une adhésion plus prononcée du locuteur à la vérité de la complétive que la locution *avoir l'impression*. A partir de ces réflexions, les verbes pris en compte par Borillo (1982) sont classifiés à la fois selon des critères syntaxiques et sémantiques, comme nous pouvons l'observer dans la Figure 2 :

1	2	3	
<i>je crois</i>	<i>c'est probable</i>	<i>je sais</i>	<i>c'est évident</i>
<i>je trouve</i>	<i>je suppose</i>	<i>je me rends</i>	<i>je suis sûr</i>
<i>je pense</i>	<i>je présume</i>	<i>compte</i>	<i>je suis certain</i>
etc.	<i>j'espère</i>	<i>je vois</i>	etc.
	etc.	etc.	
(v. transpa- rents)		(v. semi-factifs)	
assertifs faibles		assertifs forts	

Figure 2 : classification des verbes assertifs faibles et assertifs forts de Borillo (1982).

Tous les verbes qui, dans leur emploi à la première personne du singulier du présent de l'indicatif, à la forme déclarative affirmative, « traduisent soit de manière explicite, soit de manière inférentielle, l'adhésion que le locuteur apporte à la vérité de la proposition qu'il énonce » (Borillo 1982 : 35), sont des modalisateurs d'assertion. Les assertifs « faibles » expriment un degré minime de certitude par rapport à la vérité de la complétive qu'ils accompagnent et relèvent donc de l'épistémique ; les assertifs « forts » renvoient au contraire à la notion de connaissance.

En plus de ces réflexions d'ordre sémantique, Borillo (1982) tient compte de critères syntaxiques dans son analyse des modalisateurs d'assertion : sur la base de leur propriété syntaxique à se coupler ou non à la négation dans les formules de réponse du type *Je V*, ces prédicats sont classés dans trois groupes différents. Le premier groupe comprend les verbes *croire*, *penser* et *trouver*, qui peuvent être utilisés dans leur forme positive ou négative comme réponse aux interrogatives polaires, comme dans l'exemple suivant :

(13) Est-ce qu'il est rentré ?

Je crois.

Je ne crois pas.

(Borillo 1982 : 37)

Le deuxième groupe contient des verbes, comme *supposer* ou *présumer*, qui ne peuvent pas être mis à la forme négative en réponse à une interrogative polaire :

(14) Est-ce qu'il est rentré ?

Je suppose.

*Je ne suppose pas.

(Borillo 1982 : 37)

Ces deux premiers groupes correspondent aux « assertifs faibles » et expriment la modalité épistémique. Observons d'ailleurs que le concept de « verbe transparent » mobilisé par Borillo pour décrire les verbes appartenant au groupe 1 (cf. Figure 2) est beaucoup plus étroit que la notion de « transparence » évoquée par Recanati (1984). En effet, Borillo réserve la notion de transparence à *je pense*, *je crois* et *je trouve* et la lie à celle d'assertivité faible, alors que pour Recanati des verbes tels que *je suppose* ou *je présume* sont tout aussi transparents pour autant qu'ils permettent de préfacer un contenu *p*.

Le troisième groupe coïncide avec les « assertifs forts » et contient les verbes qui peuvent être utilisés en réponse aux interrogatives polaires uniquement dans leur forme négative. L'exemple typique est constitué par le verbe *savoir* :

(15) Est-ce qu'il est rentré ?

Je ne sais pas.

*Je sais.

(Borillo 1982 : 37)

Au sein des verbes assertifs forts, Borillo opère une différence entre deux sous-types de verbes : les prédicats comme *je suis sûr·e*, *je suis certain·e* « expriment la conviction, en l'occurrence très ferme, qu'a le locuteur de la vérité de sa proposition » (Borillo 1982 : 45), alors que d'autres prédicats comme *je sais*, *je vois* « expriment le témoignage d'un savoir à travers la perception, le raisonnement, la mémoire » (Borillo 1982 : 46).

L'apport des analyses de Borillo (1982) est d'avoir éclairci le rapport entre verbes parenthétiques et assertion, en soulignant le rôle actif de ces verbes dans la modalisation de la complétive qu'ils accompagnent et en proposant de les classer sur une échelle de force assertive allant des prédicats renvoyant à la notion de croyance aux prédicats s'orientant vers le pôle de la connaissance.

2.2.2 La rection faible

En se focalisant sur les aspects principalement syntaxiques de certaines constructions verbales telles que *je pense, je crois (bien), je trouve, il me semble, il paraît*, Blanche-Benveniste (1989) développe une théorie nouvelle et originale qui influencera grandement les analyses ultérieures des verbes parenthétiques dans le contexte francophone. La catégorie de verbes qui l'intéresse, qu'elle appelle « recteurs faibles », est toutefois plus restreinte que la catégorie traditionnelle des verbes parenthétiques. En effet, si le terme « parenthétique » inclut traditionnellement tout verbe ayant une possibilité de double construction (que ce soit un prédicat exprimant une opinion, un sentiment, une sensation, une parole, etc.) la classe des « recteurs faibles » se distingue des verbes de discours directs.

Blanche-Benveniste (1989) met en relief, comme il a été fait pour les prédicats anglais, la particularité de certains verbes français introduisant une complétive d'avoir une double construction : soit en position initiale, soit en position finale ou médiane (à savoir entre deux constituants syntaxiques d'un même noyau prédicatif). Cette caractéristique est partagée par les verbes de discours direct, comme *dire* ou *expliquer*, et par ceux qu'elle appelle « recteurs faibles », comme *croire* ou *penser*. Les verbes « recteurs forts » sont au contraire ceux qui ne peuvent pas être utilisés à la fois comme introducteurs de complétive et dans des propositions intercalées. Pour distinguer un verbe recteur fort d'un verbe recteur faible, deux tests sont proposés par l'auteur. Le premier (test de pronominalisation), vérifie que la complétive soit régie par le verbe : si c'est le cas, le verbe est un recteur fort. Avec le deuxième (test de déplacement), la mobilité syntaxique du verbe est testée : si celui-ci est mobile, il s'agit d'un recteur faible. Nous exemplifions ces deux tests ci-dessous :

(16) Les événements n'empêchent pas qu'on parte.

(Blanche-Benveniste 1989 : 61)

Pronominalisation → Les événements n'empêchent pas cela.

Déplacement → *Qu'on parte, les événements n'empêchent pas.

(17) Je vous ai prouvé que c'était dans le journal.

(Blanche-Benveniste 1989 : 62)

Pronominalisation → Je vous l'ai prouvé.

Déplacement → *C'était dans le journal, je vous ai prouvé.

Dans (16) et (17), les verbes sont des recteurs forts : la complétive peut être reformulée par le pronom d'objet direct, mais la postposition du verbe n'est pas acceptable. Inversement, nous sommes face à deux verbes recteurs faibles dans les exemples suivants :

(18) On dirait qu'elle a des poux.

(Blanche-Benveniste 1989 : 62)

Pronominalisation → ?On le dirait. / ?On dirait cela.

Déplacement → Elle a des poux, on dirait.

(19) Il paraît qu'il est malade.

(Blanche-Benveniste 1989 : 62)

Pronominalisation → ?Il le paraît. / ?Il paraît cela.

Déplacement → Il est malade, il paraît.

La pronominalisation de (18) et (19) n'est pas acceptable mais le déplacement en postposition ne pose aucun problème pour les deux verbes : *on dirait que* et *il paraît que* sont des recteurs faibles et ne connaissent d'ailleurs que ce type d'emploi⁸. Pour les verbes tels que *sembler*, qui peuvent se construire aussi bien avec le subjonctif qu'avec l'indicatif, seul l'indicatif est compatible avec un emploi du verbe en incise, d'où l'agrammaticalité du deuxième énoncé dans (21) :

(20) Il semble qu'il est malade. → Il est malade, il semble.

(21) Il semble qu'il soit malade. → *(Qu')il soit malade, il semble.

⁸ Pour les verbes, comme *sembler*, qui peuvent se construire aussi bien avec le subjonctif qu'avec l'indicatif, seul l'indicatif est compatible avec un emploi du verbe en incise : *il est malade, il semble* vs. **(qu')il soit malade, il semble*.

Selon les contextes, certains verbes peuvent avoir un double fonctionnement de recteurs forts ou faibles. C'est le cas, par exemple, du verbe *empêcher*, dont nous avons mis en évidence l'emploi de recteur fort dans (16). Dans l'exemple suivant, *empêcher* est utilisé en tant que recteur faible :

(22) Ça n'empêche qu'on pouvait poser des questions.

(Blanche-Benveniste 1989 : 62)

Pronominalisation → ?Ça ne l'empêche pas. / ?Ça n'empêche pas cela. / *Ça ne l'empêche pas, qu'on pouvait poser des questions.

Déplacement → On pouvait poser des questions, ça n'empêche.

On remarque que dans cet exemple, le verbe *empêcher* est difficilement pronominalisable : le double marquage rend l'incompatibilité plus claire que la pronominalisation simple au moyen de *l'* ou *cela*. Ce double marquage résulte très naturel pour les verbes recteurs forts et pratiquement impossible pour les verbes recteurs faibles. Ainsi, Blanche-Benveniste (1989 : 63) arrive à la conclusion que « dans les emplois de recteurs faibles, le verbe ne régit pas vraiment la séquence qui suit. Il garde des apparences de verbe recteur, mais il est dépourvu d'une des caractéristiques essentielles des verbes recteurs forts », à savoir la possibilité de pronominalisation.

En plus de leur position mobile dans l'énoncé, Blanche-Benveniste met en avant une autre caractéristique des verbes recteurs faibles qui avait déjà en partie été abordée par Borillo (1982), à savoir le fait que si un verbe est porteur d'une modalité (par exemple assertive) il ne pourra pas en exprimer une autre (par exemple négative). Ainsi, les verbes comme *je crois bien, je pense bien, il paraît* sont fixés dans la modalité assertive et n'ont pas de correspondant négatif ; de la même manière, il existe des verbes fixés sur une modalité négative (par exemple *il n'empêche* ne peut pas être utilisé avec une modalité assertive). D'autres blocages qui caractérisent les recteurs faibles portent sur les personnes et les temps. Par exemple, le verbe *croire* dans son usage de recteur faible est majoritairement utilisé à la première personne du singulier, au présent de l'indicatif (*je crois, je crois bien*). Bien qu'il ne soit pas possible de fournir une liste complète des verbes recteurs faibles, ces verbes sont identifiables par une désémantisation et par la perte de leur pouvoir rectionnel, ce qui les rapproche des adverbes dits « prédicatifs » tels que *heureusement que, peut-être que, bien sûr que* (Blanche-Benveniste 1989 : 67), lesquels présentent également la possibilité de double construction (introduceurs de complétive et incise). Cette position sur la perte de rection sera

néanmoins nuancée par la suite par Blanche-Benveniste et Willems (2007 : 247), qui renoncent finalement à parler d'une absence totale de rection en avançant que les verbes recteurs faibles, « construits avec une séquence en *que*, en position incidente ou disjointe, conservent les mêmes sélections lexicales, donc les mêmes valences » (Blanche-Benveniste et Willems 2007 : 250). En témoigne notamment la reprise possible par des proformes comme *ce que* des certains verbes recteurs faibles, comme le montre l'exemple suivant :

(23) Ce que je trouve en gros c'est que plus c'est moderne moins on en parle.

(Blanche-Benveniste et Willems 2007 : 232)

Ainsi, une relation syntaxique entre verbe recteur faible et la séquence qui l'accompagne existerait bel et bien, même s'il ne s'agit pas d'une subordination au sens traditionnel du terme.

Un approfondissement des propriétés sémantiques et notamment des préférences de sélection lexicale d'un certain nombre de verbes recteurs faibles a été entrepris dans Blanche-Benveniste et Willems (2007) pour *je crois*, *je pense*, *je trouve* et dans Willems et Blanche-Benveniste (2010) pour *il me semble*, *il paraît*, *j'ai l'impression*, *on dirait* et *je dirais*. Dans les corpus de français parlé pris en compte, tous ces marqueurs sont utilisés par les locuteurs dans le but de limiter la validation de ce qui est affirmé, parfois au moyen d'un entassement de plusieurs recteurs faibles exprimant la précaution et l'approximation (Blanche-Benveniste et Willems 2007 : 239). Malgré les similitudes entre les fonctions discursives des recteurs faibles, chacun tend à sélectionner des contenus différents, que nous résumons ci-dessous :

- *Je crois* a une plus forte proportion d'emplois en incise ; il sélectionne moins souvent des évaluatifs que des approximations sur la quantité ; il se combine moins souvent que *je trouve* et *je pense* avec des marqueurs du non savoir comme *je sais pas* ;
- *Je pense* est utilisé moins souvent en incise et se distingue des autres car il accompagne souvent des futurs et des conditionnels ;
- *Je trouve* est le recteur faible avec la plus faible proportion d'emplois en incise ; il est employé dans 56% des cas avec des évaluatifs et il ne porte pas sur la quantification ; par rapport à *je crois* et *je pense*, il est utilisé trois fois plus avec le pronom personnel *moi* ;
- *Il me semble* et *j'ai l'impression* sont souvent associés à *je sais pas* et sont utilisés pour pointer l'absence d'indices suffisants concernant l'état de choses qui est présenté ;
- *Il paraît* indique que le locuteur n'est pas impliqué car la source du savoir est indirecte ;

- *J'ai l'impression* et *je dirais/on dirait* sont utilisés pour spéculer sur l'état de choses, avec une association fréquente de *j'ai l'impression* avec le futur et une préférence de *je dirais/on dirait* pour l'approximation par comparaison.

2.2.3 Problématiques liées au contexte

Les travaux plus récents de Gachet (2009, 2015) ouvrent de nouvelles pistes de réflexion sur le statut syntaxique des verbes parenthétiques. En particulier, l'idée plus ou moins admise que les verbes constructeurs de *que-P* sont, au même titre que les verbes utilisés en incise, un élément secondaire pour la transmission de l'information, n'est pas acceptée par Gachet (2015), qui démontre la fragilité des tests adoptés pour corroborer cette hypothèse (enchaînement, transparence à la négation et à l'interrogation, constructions imbriquées). Pour lui, la portée syntaxique du verbe est plus facilement décidable en contexte, même s'il n'est pas toujours évident, et parfois carrément pas possible, de « déterminer quelle partie d'un énoncé contient l'information principale » (Gachet 2015 : 235).

De la même manière, à un niveau sémantique, il n'est pas toujours possible de trancher clairement entre sens fort et sens faible du verbe. L'exemple suivant, décrit par Willems et Blanche-Benveniste (2010) comme présentant un emploi faible du verbe *croire*, est remis en question :

(24) Je crois que les profs n'ont pas le temps de se parler.

(Willems et Blanche-Benveniste 2010 : 566)

Bien que l'interprétation de *je crois que* comme recteur faible soit tout à fait possible (*les profs n'ont pas le temps de se parler, je crois*), il n'y a rien non plus qui puisse exclure une interprétation de *je crois que* en tant que recteur fort (*voici ma conviction : les profs n'ont pas le temps de se parler*). Même avec plus de contexte, cela peut rester indécis.

Gachet critique et nuance également le rapprochement traditionnel entre sens fort et emploi descriptif, ainsi qu'entre sens faible et emploi mitigateur. Cela est justifié par plusieurs analyses, qui montrent qu'un verbe recteur fort peut exercer une fonction mitigatrice sur l'énoncé qu'il introduit et que les recteurs forts peuvent également apparaître en incise. Prenons les deux exemples suivants :

(25) Il croit qu'il va pleuvoir.

(Gachet 2015 : 250)

- (26) C'était elle... elle avait été chercher de l'eau-de-vie et, en m'attendant, en avait bu, je crois vraiment, une bouteille... et l'extraordinaire de cela, le lendemain, point la moindre trace de la pocharderie de la veille !

(Gachet 2015 : 259)

Le verbe *il croit que* dans (25), étant conjugué à la troisième personne, n'est pas interprétable comme une pure mitigation de *p* et son emploi est traditionnellement décrit comme descriptif. Or Gachet montre que cet exemple se prête beaucoup plus facilement à une interprétation mobilisant le sens faible de *croire* (*il a l'impression qu'il va pleuvoir*), qu'à une interprétation activant le sens fort (*il a la conviction qu'il va pleuvoir*). Nous sommes ici face à « un vide théorique : on ne peut décider s'il s'agit du verbe faible ou du verbe fort. Il paraît donc malheureux de lier indissociablement l'emploi descriptif avec le sens 'fort' du verbe » (Gachet 2015 : 250). Avec l'exemple (26), c'est une autre idée reçue sur les verbes faibles qui est remise en discussion par l'auteur, à savoir le fait que seul l'emploi faible du verbe peut apparaître en incise. Dans (26), l'ajout de l'adverbe *vraiment* à *je crois* contraint l'interprétation du verbe *croire* à son sens fort, car « il exprime plutôt une conviction qu'une simple impression » (Gachet 2015 : 259)⁹. Gachet propose alors de ne pas voir l'effet mitigateur d'un verbe parenthétique comme opposé à une fonction descriptive, mais plutôt comme s'y superposant (Gachet 2015 : 252). En effet, dès que le locuteur se met en jeu dans son énoncé, même avec un recteur fort, il existe déjà un effet de mitigation par rapport à la simple énonciation de *p*. Sa conclusion est donc que lorsque le verbe est utilisé avec un sens descriptif, il garderait tout de même un effet modalisateur : cela expliquerait ainsi le sémantisme propre à ces verbes et leur non-interchangeabilité. En effet, s'ils « ne gardaient qu'une valeur mitigatrice excluant toute fonction descriptive, ils entreraient tous dans les mêmes contextes » (Gachet 2009 : 25).

Les points avancés par Gachet sont des questions tout à fait pertinentes qu'il nous paraît important de tenir en compte. C'est en raison de ces problématiques que notre étude n'ambitionne pas de distinguer sens fort ou faible des verbes objet de cette thèse, mais qu'elle étudie pour chacun leur environnement lexical, afin que celui-ci puisse aider à comprendre de façon approfondie les propriétés sémantiques propres à chacun d'eux. En effet, comme on l'a vu, ces aspects ont très souvent été laissés de côté dans les études sur les verbes d'indication de

⁹ Voir également le test proposé par Schneider (2007) et présenté dans la section 2.4.

point de vue¹⁰, en faveur d'une analyse plus globale d'une seule et unique classe de prédicats aux propriétés mitigatrices.

2.2.4 Approches pragmatiques

Les études plus récentes sur les verbes parenthétiques privilégient une approche pragmatique en attribuant à ces structures des fonctions de marqueurs de discours. La thèse de la grammaticalisation des verbes parenthétiques avancée pour l'anglais (cf. section 2.1.2) est partagée par Apothéloz (2003) et Andersen (2007) en ce qui concerne les trois verbes recteurs faibles français les plus communs : *je pense, je crois, je trouve*. Apothéloz reprend et complète les observations de Blanche-Benveniste (1989) sur les verbes recteurs faibles et relève les propriétés suivantes :

- L'impossibilité de pronominalisation de la complétive ;
- L'érosion morphophonologique du *que* à l'oral, attestée notamment dans certaines variétés de français parlées au Québec et en Ontario. Celle-ci se manifeste soit par « disparition pure et simple », soit par « agglutination à la forme fléchie du verbe » (Apothéloz 2003 : 244) ;
- L'insensibilité à la négation et à l'interrogation. Une négation ou une interrogation qui portent sur un verbe recteur fort affectent sémantiquement ce dernier uniquement, sans remettre en cause le contenu propositionnel de la complétive ; inversement, avec un verbe recteur faible nié ou soumis à interrogation, c'est *p* qui est nié ou interrogé.
- La fixation sur une personne et un temps grammatical (très systématique pour la première personne du présent de l'indicatif) ;
- L'a-topicalité, car la rection faible « déplace le centre de gravité topical sur la complétive » (Apothéloz 2003 : 248). Cela signifie que dans un énoncé où le verbe qui apparaît est un recteur faible, l'énoncé est informationnellement orienté vers *p*. Ainsi, la rection faible implique que le verbe « perde son sens plein au profit d'une valeur modale, [...] paraphrasable par *à mon avis* » (Apothéloz 2003 : 248) et qu'il perde également toute force assertive, qui se reporte sur la complétive.

¹⁰ Nous présentons dans la section 2.2.5 les études qui ont le mieux creusé ces différences et qui ont effectué des descriptions fines permettant de mieux comprendre les contextes d'apparition des verbes sémantiquement très proches.

Ces caractéristiques, qui sont partagées dans les grandes lignes par Andersen (2007)¹¹, font en sorte que ces verbes acquièrent un statut d'« adverboïde modal » (Apothéloz 2003 : 249) et se grammaticalisent dans un emploi que Andersen (2007) appelle de « marqueur discursif propositionnel » (MDP). Les verbes recteurs faibles seraient alors très proches d'autres syntagmes adverbiaux, « eux aussi superficiellement recteurs » (Apothéloz 2003 : 250) et au contenu épistémique, tels que *certainement que, sûrement que, bien sûr que, probablement que, peut-être que*¹². Pour Andersen (2007), certains marqueurs à la deuxième personne, comme *tu sais/vous savez* ou *tu vois/vous voyez*, ont également subi un processus de grammaticalisation et sont devenus des véritables marqueurs discursifs, bien qu'avec des fonctions différentes par rapport à ceux à la première personne : les MDP à la première personne sont des marqueurs de distance, d'engagement, d'évidentialité, alors que les MDP à la deuxième personne sont des marqueurs d'interaction et d'appel à l'interlocuteur (Andersen 2007 : 14).

2.2.5 Approches sémantiques

Les verbes objets de cette étude ont été largement analysés dans la tradition logico-philosophique française, plus particulièrement dans un cadre de logique modale et des attitudes épistémiques : on s'y réfère, depuis les travaux de Russel (1985), par l'appellation « verbes d'attitude propositionnelle », mais aussi par « verbes épistémiques » ou, plus généralement, « verbes d'opinion ». Comme l'explique Gosselin (2014),

la logique épistémique s'est donné pour objectif de rendre compte, au moyen de principes généraux et explicites, de l'ensemble des inférences valides que l'on peut obtenir à partir de l'emploi des prédicats épistémiques et, par cette voie logico-référentielle, des différences sémantiques entre ces prédicats (Gosselin 2014 : 63-64).

¹¹ Andersen (2007) évoque les caractéristiques suivantes : invariabilité morphologique, optionalité sur le plan syntaxique, position libre à l'intérieur de l'énoncé, pas de contribution au contenu propositionnel de l'énoncé ni de modification de sa valeur de vérité, sens subjectifs ou intersubjectif.

¹² Une confirmation suggestive du rapport qui existe entre adverbes modaux et rection faible est évoquée par le fait que certaines langues créoles ont recours à un adverbe épistémique signifiant « peut-être » et prononcé [atiN], qui dériverait de l'expression anglaise *I think* (Apothéloz 2003 : 250).

Dans ce cadre, nous pouvons citer les travaux de Martin (1987, 1988), qui font recours à la notion des univers de croyance pour approfondir la différence entre *savoir que p* et *croire que p*, ainsi que celle entre *croire que p* et *penser que p* ; citons également l'étude de Guerry et al. (1993), qui inclut une dimension expérimentale, et la plus récente analyse de Bernard (2015), qui s'intéresse à ces verbes en prenant en compte l'analyse du discours. Carel (2011b), en s'inspirant de la logique de Port Royal et en prolongeant les réflexions déjà abordées dans Carel et Ducrot (2009), s'intéresse aux verbes *dire*, *penser*, *prétendre* et *savoir* et les analyse au moyen de la théorie argumentative de la polyphonie (TAP). Dans leur structure « X + V que q », les verbes d'attitude peuvent couvrir deux emplois différents : un emploi « modal » et un emploi « attributif » (Carel 2011b : 349). Lorsque l'emploi est modal, le locuteur « pose q » (Carel 2011b : 349), c'est-à-dire qu'il présente d'une certaine façon le contenu propositionnel de la complétive : c'est une manière de « moduler son engagement vis-à-vis du contenu » (Carel 2011b : 356). Avec un emploi attributif, le locuteur peut soit décrire X en lui attribuant la propriété d'avoir l'opinion q, soit décrire q en lui attribuant la propriété d'être l'opinion de X. L'emploi attributif ne correspond pas à une façon de présenter q, mais plutôt à une façon de transmettre la signification argumentative du contenu lui-même. Remarquons comment, même dans des cadres théoriques assez différents, les problématiques abordées se croisent : la double lecture modale-attributive proposée par Carel (2011b) rejoint les problématiques abordées jusqu'à présent concernant le rôle sémantico-pragmatique des verbes parenthétiques. En particulier, la lecture modale semble correspondre aux emplois où le verbe parenthétique fonctionne comme un marqueur de mitigation.

Nous allons à présent nous focaliser sur un certain nombre d'analyses qui ont le mérite d'avoir contesté la conception unitaire du rôle sémantique et discursif de certains verbes d'indication de point de vue, en mettant en relief des différences sémantiques fine entre prédicats très proches sémantiquement.

2.2.5.1 Je pense que, je crois que, je trouve que

Les études que nous présentons dans cette section se sont focalisées sur les trois verbes d'indication de point de vue les plus fréquents, à savoir *je pense que*, *je crois que* et *je trouve que*. Ducrot (1975) et Gosselin (2015) ont notamment expliqué le comportement très spécifique de *je trouve que* à l'intérieur d'un paradigme plus large d'autres verbes d'indication du point de vue.

Prédication originelle et prédication seconde

L'analyse de Ducrot (1975) est particulièrement pertinente et innovative car elle fournit une description sémantique très approfondie et précise de l'expression *je trouve que*, en s'intéressant à ses différentes conditions d'emploi. Un des traits qui la caractérisent est le fait qu'elle soit utilisée pour introduire une évaluation et non une situation :

Pour comprendre un grand nombre d'énoncés commençant par *Je trouve que*, [...] il faut distinguer en eux une allusion à une certaine situation (réelle ou imaginaire) et une évaluation (qualification, appréciation) relativement à cette situation. Ce qui est posé par l'énoncé, ce qui est « trouvé », ce n'est pas la situation elle-même, mais le fait qu'elle mérite l'évaluation en question. (Ducrot 1975 : 74)

Afin d'exemplifier cette observation, prenons ces énoncés :

(27) Je trouve que Jacques aura de la chance s'il est élu.

(28) *Je trouve que Jacques sera élu.

(Ducrot 1975 : 74)

L'exemple (27) présente en même temps une situation (la possible élection de Jacques) et son évaluation (*Jacques aura de la chance*), qui est introduite par *je trouve que*. La non-acceptabilité de (28) dérive du fait que *je trouve que* introduit une situation et non une évaluation.

En partant de ces premières réflexions, Ducrot (1975) arrive à mettre en lumière les caractères propres qui distinguent *je trouve que* d'autres expressions telles que *je pense que*, *je crois que* ou encore *je suis sûr-e que*. Pour expliquer cette différence, les concepts de « prédication originelle » et de « prédication seconde » (Ducrot 1975 : 80) sont employés. Observons :

(29) Je trouve que cette voiture est confortable.

(30) *Je trouve que cette voiture est une Citroën.

Dans (29), le locuteur prend l'initiative d'attribuer un prédicat nouveau (*confortable*) à un objet (*cette voiture*) et revendique la responsabilité de cet étiquetage au moyen de *je trouve que* : la prédication est originelle. L'incompatibilité, dans (30), de *je trouve que* avec *cette voiture est une Citroën* est due au fait que l'étiquette *Citroën* n'est pas nouvelle, mais elle est déjà acquise et ne vient pas du locuteur lui-même. Néanmoins, le même énoncé serait

acceptable avec un autre verbe tel que *penser*, *croire* ou *être sûr-e*, et on serait alors face à une prédication seconde :

(31) Je pense/crois/suis sûr-e que cette voiture est une Citroën.

Dans cet exemple, le locuteur ne prend pas la responsabilité de l'étiquetage (*Citroën*), mais signale simplement que l'objet (*cette voiture*) pourrait se trouver en possession de cette propriété. Ainsi, dans une prédication seconde, « les jugements préalables sont considérés comme acquis, et [...] on se fonde sur eux sans en faire l'objet même de l'activité de parole » (Ducrot 1975 : 79). En somme, pour qu'un énoncé avec *je trouve que* soit acceptable, il faut que « l'affirmation de *p* soit présentée comme l'attribution originelle d'une certaine qualité à un certain objet et ne serve pas seulement à signaler que l'objet se trouve être possesseur de cette qualité » (Ducrot 1975 : 80).

Ducrot (1975 : 84) conclut son analyse au moyen d'un tableau récapitulatif exposant les propriétés de plusieurs de ces constructions verbales. Ce tableau permet de mettre en perspective l'emploi de *je trouve que* parmi une classe plus générale de verbes :

<i>Considérer</i>	+ P	+ M	+ O	+ C	+ R
<i>Trouver</i>	+ P	+ M	+ O	- C	- R
<i>Estimer</i>	+ P	+ M	- O	+ C	+ R
<i>Juger</i>	+ P	- M	- O	+ C	+ R
<i>Avoir l'impression</i>	+ P	- M	- O	- C	- R
<i>Etre sûr</i>	- P	- M	- O	+ C	- R
<i>Penser</i>	- P	- M	- O	- C	+ R
<i>Croire</i>	- P	- M	- O	- C	- R

Figure 3: classification de Ducrot (1975).

Les lettres P, M, O, C et R représentent cinq critères différents, que nous expliquons ci-dessous ; les symboles + et - signalent si le verbe se trouve ou non en possession dudit critère. Les verbes sont listés à leur forme à l'infinitif, mais les emplois analysés sont ceux à la première personne du présent de l'indicatif.

- Critère P : le jugement du locuteur est fondé sur une expérience directe ou indirecte. Par exemple, il serait trompeur pour le locuteur d'affirmer *je considère que ce film est intéressant* s'il n'a pas vu le film en question ou si personne ne le lui a raconté (Ducrot 1975 : 75). Les expressions qui possèdent cette propriété sont *je considère que*, *je trouve que*, *j'estime que*, *je juge que*, *j'ai l'impression que*.
- Critère M : le locuteur a une expérience de la chose elle-même. Cela signifie que le locuteur exprime un jugement intrinsèque, « fondé sur la considération de l'objet même qui est jugé » (Ducrot 1975 : 77). Afin d'exemplifier ce critère, Ducrot revient

sur l'exemple du film : le fait que le locuteur puisse affirmer *j'ai l'impression que ce film est intéressant* est dû au fait qu'il n'a pas besoin de connaître le film lui-même. En effet, « il peut suffire de savoir qu'il a eu du succès [...] ou de connaître son metteur en scène, ses acteurs principaux, son thème général, etc. » (Ducrot 1975 : 77). Les critères M et P convergent en partie, car lorsque le locuteur a une expérience de la chose elle-même (critère M), son jugement est forcément fondé sur une expérience (critère P). Observons en effet que les verbes qui possèdent le critère M (*je considère que, je trouve que, j'estime que*) possèdent également systématiquement le critère P.

- Critère O : le verbe est uniquement compatible avec une prédication originelle (*je trouve que, je considère que*).
- Critère C : le locuteur se présente comme certain de l'opinion exprimée dans la complétive (*je considère que, j'estime que, je juge que, je suis sûr que*).
- Critère R : le locuteur présente son opinion comme le produit d'une réflexion (*je considère que, j'estime que, je juge que, je pense que*).

Parmi les verbes que nous prenons en compte dans ce travail, nous observons que :

- Seuls *je trouve que* et *je considère que* ne sont pas compatibles avec une prédication seconde ;
- *Je pense que* et *je crois que* ont des conditions d'emploi presque superposables. Le seul point de différence est représenté par le critère R. Cela signifie que pour Ducrot (1975) le locuteur présente son contenu comme le fruit d'une réflexion avec *je pense que* mais jamais avec *je crois que*.
- *Je suis sûr que* a un profil qui se rapproche beaucoup plus de *je pense/crois que* que de *je trouve/considère que*.

Développement ultérieurs

L'étude de Gosselin (2015) poursuit les réflexions de Ducrot (1975) et exemplifie le fonctionnement de certains verbes d'indication de point de vue en faisant recours à la Théorie Modulaire des Modalités (Gosselin 2010). Pour lui, ce qui distingue *je trouve que* d'autres expressions telles que *je pense que* ou *je crois que* est le fait que *je trouve que* n'est pas compatible avec les prédicats aléthiques, à savoir ceux qui sont présentés comme vrais « indépendamment de l'auteur du jugement » (Gosselin 2015 : 35). Prenons les exemples suivants :

(32) *Je trouve que cette table est rectangulaire.

(33) Je trouve que ce pain est bon.

(34) Je trouve que cet homme est malhonnête.

(Gosselin 2015 : 37)

Je trouve que est compatible avec des modalités appréciatives (comme celle exprimée par l'adjectif *bon* dans l'exemple (33), au moyen duquel le locuteur donne son avis sur le goût du pain) et avec des modalités axiologiques (comme celle induite par l'adjectif *malhonnête* dans (34), qui implique un jugement moral de la part du locuteur). Observons que lorsque le locuteur a recours à une modalité appréciative ou axiologique, il effectue une prédication originelle, car il colle une nouvelle étiquette (dans les exemples ci-dessus *bon* et *malhonnête*) aux prédicats. L'exemple (32) montre l'incompatibilité de *je trouve que* avec les modalités aléthiques, à savoir celles qui expriment des jugements objectifs, qui sont vrais indépendamment du locuteur qui les énonce. Nous retrouvons ainsi dans cet exemple une prédication seconde, car le concept de *rectangulaire* est toujours le même, il est partagé et acquis et sa perception ne varie pas de locuteur en locuteur.

Contrairement à *je trouve que*, *je pense/crois que* sont compatibles avec tout type de prédicat (aléthique, épistémique, appréciatif et axiologique). Lorsqu'ils sont utilisés avec des prédicats aléthiques, ils ont une valeur atténuative. Avec les prédicats épistémiques, appréciatifs et certains prédicats axiologiques, seul *je pense que* peut pour Gosselin avoir un double fonctionnement : il peut soit atténuer *p*, soit être utilisé comme un « marqueur de subjectivité individuelle » (Gosselin 2015 : 41). Il viendrait alors préciser le caractère individuel du point de vue du locuteur en se rapprochant sémantiquement de *je trouve/estime/considère que*. Prenons l'énoncé suivant :

(35) Je pense que cette décision est juste.

(Gosselin 2015 : 41)

Dans cet énoncé, *je pense que* introduit un jugement axiologique qui se manifeste par la présence de l'adjectif *juste*. Selon l'analyse proposée par Gosselin (2015), cet énoncé peut avoir une double lecture : soit *je pense que* atténue l'assertion *p* (*cette décision est juste*), introduisant une certaine réserve, soit cette construction fonctionne comme un marqueur de subjectivité individuelle qui souligne que le jugement relève de l'opinion du locuteur. Si on met en lien cette analyse avec celle de Ducrot (1975), pour qui les constructions telles que *je pense que* et *je crois que* peuvent s'utiliser avec une prédication originelle ou avec une

prédication seconde, on peut observer que la double interprétation est dépendante de l'interprétation de la prédication elle-même. En effet, si la prédication est seconde (l'étiquetage est déjà accompli et partagé), *je pense que* aura une fonction d'atténuateur de *p* ; si la prédication est originelle (l'étiquetage est nouveau), *je pense que* aura une fonction de marqueur de subjectivité individuelle.

Si *je pense que* est, pour Gosselin, le seul verbe à pouvoir fonctionner et comme atténuateur et comme marqueur de subjectivité individuelle, pour Rendulić et Kanaan-Caillol (2016) *je crois que* peut tout aussi bien remplir ces deux fonctions. Angot (2021 : 95) penche également pour cette dernière option, en attribuant à *je pense* et *je crois*, selon le contexte, un statut de marqueurs épistémiques (« epistemic marker ») ou un statut de marqueurs de subjectivité (« subjective marker »).

2.2.5.2 *Je considère que*

L'expression *je considère que* a également fait l'objet d'études sémantiques, qui ont essayé de mettre en lumière ce qui la distingue de *je trouve que* et de *j'estime que* (cf. Gosselin 2015, Tuchais 2012). Pour Gosselin (2015), *je considère que* et *j'estime que* se rapprochent sémantiquement de *je trouve que* car tous précisent la valeur subjective de la modalité intrinsèque et n'ont donc pas de valeur atténuative. *J'estime/considère que* sont donc, comme *je trouve que*, difficilement compatibles avec des prédicats aléthiques. Néanmoins, des différences sémantiques peuvent être mises en relief : si *je trouve que* est compatible avec une « subjectivité individuelle temporellement variable » (Gosselin 2015 : 39), *je considère que* et *j'estime que* requièrent une « certaine stabilité temporelle du jugement » (Gosselin 2015 : 39). Observons ces deux exemples :

(36) Cette semaine, je suis grippé, et je trouve que le pain est bon et que le fauteuil est inconfortable.

(37) Cette semaine, je suis grippé, et je ??considère/ ??estime que le pain est bon et que le fauteuil est inconfortable.

Gosselin (2015: 39)

Dans (36), le locuteur exprime, au moyen de *je trouve que*, des sensations qui sont dues à son état (passager) de malaise ; l'acceptabilité de *je considère/estime que* dans (37) est plus difficile étant donné le caractère non adapté de ces deux verbes, qui impliquent un jugement ancré dans la temporalité, à décrire une situation qui ne l'est pas. On peut également avancer

que cette incompatibilité est due au fait que ces deux verbes impliquent « un minimum d'activité réflexive » (Tuchais 2012 : 45), ce qui serait inapproprié pour décrire de simples sensations. Cette propriété est en ligne avec l'analyse de Ducrot (1975). Rappelons que *je considère que* et *j'estime que* sont tous deux porteurs des critères C et R (qui impliquent respectivement que le locuteur se présente comme sûr du point de vue exprimé dans *p* et que ce point de vue est le fruit d'une réflexion), alors que *je trouve que* ne l'est pas.

2.2.5.3 Et les autres verbes du paradigme ?

Des études spécifiques aux autres verbes d'indication de point de vue de notre paradigme manquent, notamment pour ce qui concerne *je dis que* et *je suis convaincu-e que*. Ces expressions n'ont fait ni l'objet de descriptions sémantiques approfondies concernant leur emploi à la première personne du singulier en tant qu'expressions introductrices de complétive, ni été étudiées dans le cadre de leur emploi en incise.

Le verbe *sembler*, quant à lui, est majoritairement étudié par opposition au verbe *paraître*, dans les constructions impersonnelles *il semble que* et *il paraît que* (cf. notamment les études de Nølke 1994 et Carel 2011b). Bien qu'interchangeables dans certains contextes, ces deux expressions ne le sont pas systématiquement et ont ainsi un sémantisme différent. En effet, *il paraît que* est un « marqueur évidentiel du type oui-dire » (Nølke 1994 : 86), alors que *il semble que* implique une idée « d'atténuation ou d'hésitation » du locuteur par rapport à *p*. Lorsque le locuteur ajoute le pronom *me* à *il semble que*, ses réserves diminuent et sa responsabilité par rapport à la vérité de ses propos augmente : si « l'adjonction du datif *me* diminue l'effet de distance [...], c'est parce que ce pronom indique que le locuteur prend quand même la vérité de *p* à son compte – après coup » (Nølke 1994 : 93). Willems et Blanche-Benveniste (2009) et Willems (2011) s'alignent à cette analyse et catégorisent *il me semble (que)* en tant que verbe recteur faible : le locuteur prend en charge *p* tout en mettant en avant l'absence d'indices suffisants, ce qui mène à la transmission d'une connaissance floue ou incertaine.

2.3 Les verbes d'indication de point de vue italiens

Contrairement aux domaines anglophone et francophone, les études sur les verbes d'introduction de point de vue en italien demeurent limitées. Cette section examine les études de Venier (1991), de Strik Lievers (2012) et de Lo Baido (2024), en mettant en lumière les

aspects les plus pertinents pour notre étude. Enfin, elle aborde également quelques questions plus spécifiquement syntaxiques.

2.3.1 Venier (1991)

L'étude la plus approfondie sur les verbes d'indication de point de vue italiens est la monographie de Venier (1991), qui s'intéresse aux verbes parenthétiques *credo* et *penso* en abordant trois de leurs caractéristiques principales : la question de leur non-négabilité, celle de leur rapport avec la performativité et celle concernant le degré de subjectivité que ces deux verbes expriment.

Venier s'insère dans la tradition qui rapproche les verbes parenthétiques des adverbes modaux à cause de leur rôle discursif très proche. En particulier, elle confirme l'hypothèse d'Urmson (1952 : 487) selon laquelle les verbes parenthétiques exprimeraient un degré de subjectivité supérieur par rapport aux adverbes modaux. En effet, ces deux éléments caractérisent de façon différente le processus inférentiel de l'assertion dans laquelle elles apparaissent. L'adverbe modal présente le processus inférentiel comme conclu, tout en présentant les résultats de ce dernier. Le verbe parenthétique, au contraire, laisse entendre que le processus inférentiel est encore en cours. Observons :

(38) Il garage è ancora vuoto. Evidentemente i Rossi sono ancora in vacanza.

(Venier 1991: 117)

Le garage est encore vide. Manifestement, les Rossi sont toujours en vacances.

(39) Il garage è ancora vuoto. Deduco che i Rossi sono ancora in vacanza.

(Venier 1991: 117)

Le garage est encore vide. J'en déduis que les Rossi sont toujours en vacances.

Venier (1991) postule qu'au moyen d'un adverbe modal, comme dans (38), l'interlocuteur est amené à reconstruire un contexte qui justifie l'assertion, alors qu'avec un verbe parenthétique (39) il accède uniquement au contexte personnel du locuteur et à son évaluation de la situation.

2.3.2 Cignetti (2009)

L'étude de Cignetti (2009) adopte une méthodologie basée sur corpus pour analyser la sémantique du verbe *credo*. Il constate tout d'abord que *credo* est plus fréquent dans le corpus

oral C-ORAL ROM (88,2%) que dans le corpus écrit LISUB (11,8%) et qu'il apparaît davantage dans des contextes formels que dans des contextes informels. L'hypothèse pour cette tendance est qu'une plus grande distance entre les interlocuteurs dans les contextes formels entraîne une utilisation plus fréquente de formes de mitigation, dont *credo* fait partie. En effet, dans certains contextes, *credo* est utilisé comme un écran – « uno schermo » (Cignetti 2009 : 943) – que le locuteur utilise pour se protéger par avance des contre-arguments potentiels de son interlocuteur.

Sur le plan syntaxique, *credo* a une plus grande liberté de structure à l'oral, où il peut apparaître en début, au milieu ou en fin de phrase ; en revanche, la séquence *credo che p* couvre le 96,1% des emplois à l'écrit. Si pour Urmson (1952), les trois différentes positions des verbes parenthétiques remplissent la même fonction sémantique, pour Cignetti (2009) ce n'est pas le cas. En effet, en incise ou position finale, *credo* couvre principalement une fonction épistémique, alors qu'en position initiale il a plutôt une fonction argumentative qui signale l'introduction du *claim* du locuteur (Cignetti 2009 : 949)¹³. Le fait que *credo che p* soit très souvent précédé ou suivi par des connecteurs comme *quindi* et *dunque* est pour Cignetti une preuve de cette fonction argumentative. Pour distinguer ces deux fonctions, deux tests sont proposés. Lorsque *credo* a une valeur épistémique, il peut être remplacé par *se non mi sbaglio* (si je ne me trompe pas) ou par *ma non vorrei sbagliare* (mais je ne voudrais pas me tromper) ; s'il a une valeur argumentative, la paraphrase serait *la mia opinione è che* ou *la mia tesi è che* (mon opinion/avis est que).

2.3.3 Strik Lievers (2012)

L'analyse sémantico-lexicale de Strik Lievers (2012) sur les verbes à complément prédicatif (*verbi con complemento predicativo*) mérite également d'être mentionnée dans le cadre de cette thèse, car elle prend en compte des verbes appelés *modalizzanti* (« modalisants ») tels que *credere*, *ritenere*, *giudicare* et *considerare*. Basée sur le corpus « la Repubblica » (Baroni et al. 2004), cette étude examine les adjectifs postverbaux et démontre que ces prédicats se différencient sémantiquement. Par exemple, le verbe *credere* exprime de façon prototypique un jugement épistémique car il s'accompagne le plus fréquemment d'adjectifs épistémiques

¹³ Ces deux fonctions correspondent aux emplois mitigateur-descriptif de Benveniste, ainsi qu'aux emplois modal-attributif de Carel (2011b).

comme *possibile* (possible) ou *impossibile* (impossible). Ces adjectifs n'apparaissent pas fréquemment avec le verbe *considerare* et jamais avec le verbe *giudicare* : ceux-ci ont en revanche tendance à introduire des adjectifs évaluatifs (*importante, pericoloso, positivo, indispensabile, etc.*). Quant à *ritenere*, il accompagne aussi bien des adjectifs épistémiques qu'évaluatifs, mais les plus fréquents sont *opportuno* (opportun) et *necessario* (nécessaire). Enfin, les verbes *sembrare* et *parere*, complètement interchangeables au niveau de leur sémantisme, se distinguent en ce qui concerne leurs fréquences d'usage et leurs contextes d'utilisation : dans le corpus pris en compte, *sembrare* est beaucoup plus fréquent que *parere*, qui apparaît préférentiellement dans des contextes de registre formel ou littéraire.

2.3.4 Lo Baido (2024)

Dans une approche plus pragmatique, Lo Baido (2024) s'intéresse à celles qu'elle appelle les *comment clauses*, à savoir des expressions comme *penso, credo, dico, guarda, sai*, qui servent au locuteur à se situer en fonction de critères épistémiques, rhétoriques, textuels et interpersonnels. En se basant sur le corpus KIParla (Goria et Mauri 2018 ; Mauri et al. 2019), elle relève trois types de *comment clauses* : (i) celles qui concernent l'attitude du locuteur et ses sources d'information, (ii) celles qui travaillent au niveau du domaine textuel, visant à assurer la cohérence et la cohésion textuelles, (iii) celles qui travaillent au niveau du domaine interpersonnel (qui visent par exemple à capturer l'attention de l'interlocuteur ou à chercher sa validation).

Les expressions qui nous intéressent tout particulièrement sont celles du premier type, car elles concernent le locuteur et son positionnement dans l'énoncé. Ce positionnement se manifeste principalement par une posture épistémique – au moyen de marqueurs tels que *credo, (io) penso, immagino, mi pare, mi/a me sembra, mi sa, suppongo* – et par une posture évidentielle – parmi les marqueurs évidentiels nous retrouvons *penso* et *credo*, qui sont très polyfonctionnels. Une autre fonction abordée est celle de l'emphase, qui concerne notamment le marqueur *dico*. Dans certains contextes, *dico* véhicule une nuance de désaccord : dans ce cas, d'autres marqueurs précis interviennent de façon récurrente, en particulier les adverbes *davvero, proprio, abbastanza, piuttosto*, le connecteur *ma* et la particule de négation *no* (Lo Baido 204 : 137).

En se basant sur le même corpus, Battaglia et Lo Baido (sous presse) approfondissent les propriétés syntaxiques et la distribution cotextuelle des verbes *penso, credo, immagino* et *suppongo*, afin d'évaluer dans quelle mesure ces formes participent à l'émergence de la

posture du locuteur dans son discours. Des tendances d'usage sont mises en relief : *penso* apparaît fréquemment dans des contextes évidentiels où le locuteur apporte des arguments et des preuves de ce qu'il affirme ; *credo* semble plus enclin à s'insérer dans des contextes purement épistémiques ; *immagino* et *suppongo* apparaissent de façon prototypique en position parenthétique, ce qui amène les autrices à faire l'hypothèse que leur niveau de pragmatization serait plus avancé que celui de *penso/credo*.

En conclusion, les études de Lo Baido (2024) et de Battaglia et Lo Baido (sous presse) mettent en lumière le fait que les fonctions des formes verbales prises en compte ne sont pas nécessairement systématiquement encodées dans leur sémantique lexicale. En effet, elles peuvent parfois être déclenchées par des associations récurrentes avec d'autres marqueurs. Par conséquent, le contexte joue également un rôle déterminant dans la mise en scène de ces fonctions.

2.3.5 Etudes syntaxiques

Dans une perspective syntaxique, Giorgi et Pianesi (2005) s'intéressent aux propriétés des verbes *credo* et *dicono* et en particulier au phénomène du *complementizer deletion*, à savoir le fait que la conjonction *che* peut tomber dans certains cas de figure. En règle générale, le *che* peut être omis s'il introduit une proposition présentant un verbe conjugué au subjonctif (cf. Giorgi et Pianesi 2005 : 107). Contrairement au français, ce phénomène est assez systématique en italien et en anglais (cf. 2.1.2), bien que les conditions d'omission du *che* et du *that* ne soient pas les mêmes. L'alternance libre entre indicatif et subjonctif dans la complétive introduite par un verbe comme *pensare* ou *credere* suscite également beaucoup d'intérêt (cf. Bronzi 1977, Sabatini 1985, Schena et al. 2000). Par exemple, Schneider (1999) examine l'utilisation du subjonctif après des verbes comme *credere*, *parere*, *pensare* et *sembrare* dans le corpus oral CLIP (*Corpus del lessico di frequenza dell'italiano parlato*). Cette question a été également évoquée dans une perspective comparative avec le français (cf. Soliman 2010, Begioni 2013). La particularité de l'italien est qu'il permet de plus en plus un remplacement du subjonctif par l'indicatif non seulement à l'oral, mais aussi à l'écrit (Soliman 2010 : 70). Ce procédé ne servirait pas un renforcement de la validation de *p*, en accord avec le caractère « véridique » du mode indicatif (cf. Brugger et D'Angelo 1995), mais serait plutôt utilisé comme une alternative stylistique (Soliman 2010 : 68).

2.4 Approches comparatives

A ce jour, il existe peu d'études comparatives sur les verbes d'indication de point de vue. Parmi celles-là, nous pouvons citer la monographie de Doro-Megy (2008), qui examine les verbes *penser, croire, think, believe* dans une approche comparative français-anglais ; l'étude basée sur corpus de Haßler (2014), qui s'intéresse à l'évolution diachronique de *je crois* dans trois langues romanes (français, espagnol, portugais) ; les études des Tuchais (2014, 2020) sur les *je pense, je crois* et leurs correspondants japonais.

Dans cette section, nous nous attardons particulièrement sur la monographie de Schneider (2007), car elle nous intéresse pour son analyse comparative des verbes d'indication de point de vue en français et en italien, bien que son intérêt porte sur leur usage en incise. En s'appuyant sur 22 corpus oraux en français, italien et espagnol, Schneider sélectionne plus de 80 *Reduced Parenthetical Clauses* (RPCs), qu'il analyse en s'intéressant de manière approfondie à leur rôle communicatif. Pour qu'une RPC soit retenue dans le paradigme, elle doit répondre aux sept critères suivants :

- La seule restriction morphologique est que la RPC doit être basée sur un verbe fini. En effet, en ne sélectionnant que les verbes utilisés à la première personne du singulier du présent de l'indicatif, des expressions comme *on va dire, ho detto* ou *digamos* n'auraient pas pu être repérées ;
- Il doit y avoir une absence de lien manifeste entre la RPC et la proposition hôte ;
- La RPC doit pouvoir se placer en position médiane ou finale ;
- La RPC doit avoir le pouvoir d'interrompre un lien syntaxique étroit, comme par exemple celui qui existe entre sujet et prédication ou entre verbe et complément ;
- La proposition principale est indépendante de la RPC : en ce sens, elle reste acceptable même en absence de la RPC ;
- Le verbe de la RPC doit manquer d'un argument (ainsi, *j'espère* est retenu mais *je l'espère* ne l'est pas) ;
- L'argument manquant doit pouvoir se retrouver sémantiquement dans la proposition hôte.

Le tableau suivant récapitule les RPCs qui satisfont ces critères dans les trois langues, et les accompagne des pourcentages de fréquences des RPCs dans les échantillons étudiés :

French	Italian	Spanish
<i>tu vois</i>	19.8 <i>diciamo</i>	41.0 <i>digamos</i>
<i>je veux dire</i>	11.4 <i>dico</i>	7.5 <i>digo</i>
<i>vous savez</i>	11.4 <i>credo</i>	6.4 <i>sabes</i>
<i>je crois</i>	10.2 <i>dice</i>	6.2 <i>dice</i>
<i>tu sais</i>	9.3 <i>voglio dire</i>	5.8 <i>creo PRO</i>
<i>je sais pas</i>	7.2 <i>non so</i>	5.2 <i>no sé</i>
<i>disons</i>	6.3 <i>sai</i>	3.6 <i>me parece</i>
<i>voyez-vous</i>	5.7 <i>direi</i>	3.4 <i>quiero decir</i>
<i>je dirais</i>	4.5 <i>mi sembra</i>	2.4 <i>creo</i>
<i>je pense</i>	3.9 <i>devo dire</i>	2.1 <i>ya te digo</i>
<i>j'imagine</i>	1.5 <i>penso</i>	1.9 <i>ya digo</i>
<i>mettons</i>	1.5 <i>sa</i>	1.9 <i>digo PRO</i>
<i>je trouve</i>	1.5 <i>spero</i>	1.9 <i>diríamos</i>
<i>je sais plus</i>	1.2 <i>dice NP</i>	1.7 <i>me imagino</i>
<i>je suppose</i>	1.2 <i>vedi</i>	1.5 <i>supongo</i>
<i>paraît-il</i>	0.9 <i>ho detto</i>	1.1 <i>diría</i>
<i>on va dire</i>	0.6 <i>figurati</i>	1.1 <i>insisto</i>
<i>j'espère</i>	0.6 <i>mi pare</i>	1.1 <i>me parece a mí</i>
<i>dites-vous</i>	0.3 <i>penso PRO</i>	0.9 <i>diría PRO</i>
<i>il faut dire</i>	0.3 <i>poni</i>	0.9 <i>ha dicho</i>
<i>on dirait</i>	0.3 <i>mi sa</i>	0.9 <i>me han dicho</i>
<i>je me rappelle</i>	0.3 <i>volevo dire</i>	0.6 <i>podríamos decir</i>
<i>me semble-t-il</i>	0.0 <i>vede</i>	0.4 <i>fijate</i>
	<i>immagino</i>	0.2 <i>parece</i>
	<i>sembra</i>	0.2 <i>pienso PRO</i>
		<i>repito</i>
		<i>me acuerdo</i>
		<i>decía NP</i>
		<i>dicen</i>
		<i>ha dicho NP</i>
		<i>imagino</i>
		<i>imagínate</i>
		<i>ya ves</i>

Table 4.9: Frequencies of French, Italian, and Spanish RPCs

Figure 4 : RPCs relevées par Schneider (2007 : 86).

Comme nous pouvons le constater dans la Figure 4, les RPCs construites sur le verbe *dire* sont les plus fréquentes et celles qui ont le plus grand nombre de variation de formes. En français, nous avons *je veux dire*, *disons*, *je dirais*, *on va dire*, *dites-vous*, *il faut dire*, *on dirait* et pour l'italien *diciamo*, *dico*, *dice*, *voglio dire*, *direi*, *devo dire*, *dice*, *ho detto*, *volevo dire*. Les verbes tels que *croire*, *savoir*, *penser* sont également très fréquents dans les trois langues. En français et en italien, les RPCs se retrouvent majoritairement dans des discours bidirectionnels, de type face-à-face avec un tour de parole régulé (tels que des discours dans des assemblées législatives, des examens oraux à l'école ou à l'université, des entretiens, etc.). Les RPCs sont inversement moins fréquentes dans les textes monologiques de tous domaines. Pour l'espagnol, aucune tendance nette de ce type n'a été relevée (Schneider 2007 : 89).

Au niveau pragmatique, les RPCs sont décrites comme des marqueurs de mitigation permettant au locuteur de réduire le poids de la responsabilité de ses déclarations. En d'autres mots, ce sont des *responsability alleviating devices* (Schneider 2007 : 109). Afin de décrire leurs fonctions, l'auteur se base sur la distinction qu'Hare (1970) propose entre *phrastic* (le contenu propositionnel), *tropic* (le type d'acte mis en scène, à savoir la force

illocutionnaire de l'énoncé) et *neustic* (l'implication du locuteur concernant le contenu propositionnel)¹⁴, ainsi que sur la classification de Caffi (1999, 2001) qui répartit les *mitigating devices* en *bushes*, *hedges* et *shields*. Ainsi, quatre classes principales de RPCs sont présentées :

- Les RPCs qui mitigent le *phrastic*, comme *disons/diciamo/digamos* ;
- Les RPCs qui indiquent le *tropic* et qui mitigent soit le *phrastic* soit le *neustic* : il s'agit notamment des performatifs comme *dico/digo* et les *hedged performatives* tels que *je veux dire/voglio dire/quiero decir, je dirais/direi/diría* ;
- Les RPCs qui mitigent directement le *neustic* : ce sont les parenthétiques classiques tels que *je crois* et *je pense* ;
- Les RPCs qui mitigent le *neustic* de façon indirecte, comme les verbes évidentiels *tu vois/vedi/ya ves* et les verbes de connaissance *tu sais/sai/sabes*.

L'analyse de Schneider a le mérite de montrer que la description de ces expressions en tant que marqueurs de mitigation peut être affinée en étudiant sur quel élément la mitigation porte. Néanmoins, il est parfois très difficile de mettre en évidence si la mitigation s'effectue sur le *phrastic*, sur le *tropic* ou sur le *neustic*, car certains marqueurs peuvent avoir plus qu'une fonction (Schneider 2007 : 95). Aucun test n'est d'ailleurs proposé pour distinguer entre ces différents emplois. De plus, certains marqueurs (notamment ceux de discours rapporté), peuvent avoir d'autres fonctions que celle de mitigation (Schneider 2007 : 133).

Au niveau sémantique, Schneider relie le sens lexical des prédicats des RPCs à leur fonction pragmatique : les verbes de parole (*disons, je dirais, diciamo, dico*), mitigent le *phrastic* ou le *neustic*. Les verbes d'opinion et de croyance (*je crois, je pense, credo, penso*), ainsi que les verbes d'opération mentale (*je suppose, j'imagine, je trouve, immagino*), mitigent directement le *neustic*. Enfin, les verbes de sensation et de perception (*tu vois, je me rappelle, mi pare, mi sa*) et les verbes épistémiques (*tu sais, vous savez, sa, sai*) mitigent indirectement le *neustic*.

Au niveau syntaxique, les RPCs sont des *ungoverned dependents* (Schneider 2007 : 167), à savoir des éléments qui sont dépendants de la proposition hôte, mais qui ne sont pas régis par elle – sauf quand la RPC est suivie de *que*. Afin de comprendre dans quels cas la RPC est une

¹⁴ Schneider (2007 : 92) souligne que la différence entre *tropic* et *neustic* rappelle celle de certaines grammaires françaises qui distinguent entre modalité d'énonciation (déclarative, interrogative, etc.) et modalité d'énoncé (évaluation du locuteur).

governing clause (Schneider 2007 : 191)¹⁵, l'auteur propose plusieurs critères, en plus de l'étude du contexte. Par exemple, la présence d'un adverbe venant modifier le parenthétique est un signal que le verbe a une lecture non parenthétique, comme dans l'exemple suivant :

- (40) C : [...] io credo con molta forza che sarebbe un gravissimo errore pensare di di l'area metropolitana di Napoli a se stante [...].

(Schneider 2007 : 196)

La modification de *credo* par la locution adverbiale *con molta forza* indique qu'un procédé cognitif est en cours et que la lecture est descriptive (Schneider parle de « assertive reading »). Le test du conditionnel peut également aider à distinguer les deux emplois : si le verbe peut être repris par la conjonction *si*, alors l'emploi parenthétique est exclu, car le conditionnel active le sens descriptif du verbe. Inversement, les emplois parenthétiques sont *non-thruth conditional* (voir également Infantidou (2001 : 125 ; Jayez et Rossari 2004 : 215).

¹⁵ Rappelons-nous que cette bivalence avait déjà été évoquée par Urmson (1952), qui postulait que les verbes parenthétiques peuvent avoir un « more or less pure parenthetical use » (Urmson 1952 : 481), et par Blanche-Benveniste (1989) qui distingue, pour certains verbes comme *empêcher*, entre emplois de recteur fort et emplois de recteur faible. Toutefois, aucune référence au contexte, qui pourrait aider à désambiguïser l'une ou l'autre signification, n'est faite par ces auteur·es.

3. Pour une analyse quantitative des verbes d'indication de point de vue : outils et méthode

Les études présentées dans le chapitre 2 partagent un point de départ commun : la circonscription d'un ensemble de prédicats partageant les mêmes propriétés syntaxiques (invariabilité morphologique et position libre à l'intérieur de l'énoncé dans lequel ils interviennent). Ces caractéristiques syntaxiques sont systématiquement mises en relation avec des propriétés sémantico-énonciatives, à savoir le fait que ces prédicats modifient la valeur de vérité de l'énoncé qu'ils accompagnent en lui conférant un sens modal. Bien que ces études aient permis de dégager un certain nombre de traits communs à ces verbes appelés parenthétiques ou recteurs faibles, elles ne contemplent pas la possibilité d'étude systématique de leurs contextes d'emploi. De plus, les corpus, pourtant essentiels à une telle analyse, ne sont que très rarement mobilisés. Lorsqu'ils le sont, c'est principalement pour fournir des exemples illustratifs ou pour mesurer la fréquence d'apparition de ces formes. A ce jour, nous n'avons pas connaissance d'une quelconque application de méthodes quantitatives s'intéressant à l'analyse des environnements lexicaux des verbes d'indication de point de vue.

En se concentrant sur les contextes privilégiés d'apparition de ces verbes, notre étude met en lumière des régularités qui ne pourraient être détectées par une approche purement théorique. En prenant en compte des exemples authentiques et en augmentant la taille des observables, nous disposons d'une base très riche qui permet de mettre en évidence des régularités d'emploi qui resteraient invisibles dans le cadre d'une étude fondée uniquement sur des manipulations d'exemples construits. De plus, notre méthodologie dépasse les limites d'une approche qui ne considère les corpus que comme de simples réservoirs d'exemples : elle fait émerger des tendances difficilement observables par un dépouillement manuel, une méthode chronophage et peu représentative sur le plan statistique. En combinant approche statistique et analyse qualitative, notre démarche permet ainsi de mieux cerner dans quelle mesure ces verbes apparaissent dans des contextes similaires et d'examiner d'éventuelles différences entre deux langues de même souche.

Le présent chapitre se structure de la manière suivante : dans un premier temps, nous exposons les choix méthodologiques qui ont guidé la sélection des corpus, en détaillant les étapes de leur constitution, ainsi que les outils utilisés pour les interroger (sections 3.1 et 3.2). La section 3.3 présente les paradigmes des verbes d'indication de point de vue en français et en italien, en précisant les démarches méthodologiques et théoriques adoptées pour l'extraction des deux

paradigmes de verbes. Enfin, la section 3.4 expose les fondements théoriques de l'analyse cooccurrence que nous mettons en œuvre, en décrivant son application à travers trois niveaux d'analyse distincts.

3.1 Choix des corpus

Les recherches qui ont étudié les représentations des verbes d'indication de point de vue dans les corpus se sont majoritairement appuyées sur des données orales : la spontanéité et l'informalité caractéristiques du langage parlé offrent un terrain privilégié pour l'analyse de diverses marques subjectives mises en scène par le locuteur lorsqu'il s'exprime. Mais ces manifestations subjectives ne restent pas cantonnées à l'oral, elles viennent s'intégrer tout naturellement et plus ou moins explicitement dans des textes écrits également. Dans le but de comprendre comment ces phénomènes se manifestent dans des contextes où la structuration et la planification du discours prennent une importance accrue, nous proposons d'examiner le comportement des verbes d'indication de point de vue dans des corpus relevant de l'écrit.

Il est tout d'abord nécessaire d'entreprendre une réflexion préliminaire quant au choix des corpus les plus à même de fournir un échantillon représentatif des constructions étudiées. Nous avons en effet besoin de corpus suffisamment étendus en termes de taille et particulièrement riches en occurrences de ces formes à la première personne.

À cette fin, nous avons privilégié l'analyse de corpus comprenant des textes à visée argumentative, notamment les discours présidentiels français et italiens. Ce type de discours constitue un terrain d'étude pertinent, dans la mesure où la construction de l'*ethos* du locuteur repose en grande partie sur l'emploi de marques de « subjectivité énonciative » (Maingueneau, 2002 : 64). L'objectif étant de gagner la confiance des interlocuteurs, nous postulons une fréquence significative des verbes d'indication de point de vue dans ce corpus¹⁶. Nous avons

¹⁶ Une précision mérite d'être apportée concernant le statut du discours politique. Bien que les allocutions aient pour but d'être oralisées, ce genre se rapproche beaucoup plus du code écrit que du code oral. En effet, il s'agit de textes préparés et rédigés à l'avance, qui présentent des structures formelles typiques de l'écrit. À titre d'exemple, Ursini (2007) souligne le caractère politique-institutionnel des discours présidentiels italiens, caractérisés par un haut degré de formalité et complexité, ainsi que par une densité de présentation de l'information se traduisant par une forte présence de voix lexicalisées et de mots pleins, avec un pourcentage majeur des noms par rapport aux verbes (Ursini 2007 : 202). Le caractère formel de ce genre ressort clairement dans nos corpus : citons, pour le français, les formules d'adresses telles que « Françaises, Français » et les rituels de clôture « Vive la République ! Vive la France ! ». Pour le corpus italien, nous signalons 170 occurrences du

également pris en compte des corpus contenant des textes à visée informative, plus précisément des articles de presse. Ceux-là peuvent être de nature différente : les corpus sont composés d'articles relatant des faits de manière aussi objective que possible (tels que les brèves, les comptes rendus ou les reportages), mais aussi d'articles accordant davantage de place à l'expression d'un point de vue subjectif, notamment les éditoriaux et les chroniques. C'est principalement au sein de ces derniers que nous anticipons une occurrence plus marquée des verbes d'indication de point de vue. De façon plus générale, ces constructions apparaissent fréquemment dans des articles intégrant des discours rapportés, indépendamment de la thématique abordée (actualité, politique, sport, etc.). Les journaux en ligne présentent notamment plus de discours rapportés que les journaux papiers(cf. Bonomi 2014)

Il convient enfin de souligner que notre objet d'étude porte spécifiquement sur les verbes d'indication de point de vue et que notre analyse vise à examiner leur fonctionnement au moyen des corpus sélectionnés. L'objectif n'est donc pas d'étudier les corpus en tant que tels, mais de les mobiliser comme outils permettant de mettre en lumière les spécificités des constructions verbales observées. Toutefois, nous reconnaissons que la manière dont ces phénomènes linguistiques se manifestent dépend intrinsèquement des caractéristiques des corpus considérés. Comme l'affirme Blumenthal (2008 : 35), toute observation sur une forme linguistique n'a de valeur que par rapport au corpus dans lequel elle a été effectuée :

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)

En conclusion, le choix de ces corpus – dont les caractéristiques seront détaillées dans la prochaine section – repose, d'une part, sur la nécessité de disposer d'un ensemble de données suffisamment large pour permettre une analyse approfondie des formes étudiées et, d'autre part, sur leur propension à contenir les formes linguistiques ciblées. En outre, des critères d'accessibilité et de facilité de constitution des corpus ont également joué un rôle déterminant dans cette sélection.

pronom personnel *ella* qui est utilisé à la place de *lei* : selon le dictionnaire en ligne Treccani, la forme *ella* a désormais disparu dans la langue parlée.

3.2 Constitution des corpus et outils d'analyse

Dans cette section, nous présentons de manière détaillée les corpus que nous avons interrogés. Rappelons tout d'abord qu'un corpus n'est pas un simple assemblage de données hétérogènes, mais plutôt une « collection de productions langagières attestées, écrites ou orales », qui peuvent être « des textes, des extraits de textes, des enregistrements audio ou vidéo » (Poudat et Landragin 2017 : 11). De plus, un corpus se définit par son caractère composite et sa « réflexivité », à savoir le fait que ses sous-parties renvoient les unes aux autres pour « former un réseau sémantique performant dans un tout [...] cohérent et auto-suffisant. » (Mayaffre 2002 : 6).

Pincemin (1999 : 27) met en évidence trois critères précis ou conditions à respecter par le·la linguiste lors de la constitution d'un corpus :

- I. Conditions de **signifiante** : un corpus doit être constitué en fonction d'une étude précise. Autrement dit, il doit être pertinent et cohérent par rapport aux objectifs de recherche. C'est dans cette optique que nous avons sélectionné les corpus des allocutions présidentielles et ceux de la presse, car ils s'avèrent particulièrement riches en constructions linguistiques correspondant à notre analyse¹⁷.
- II. Conditions d'**acceptabilité** : le corpus doit être le plus représentatif possible de la langue et des phénomènes étudiés, avec un niveau de détail adapté à la finesse de l'analyse qui sera menée. Comme notre étude a pour but d'accéder au sens linguistique des verbes d'indication de point de vue dans un cadre ducrotien, nous ne nous intéressons pas aux sujets parlant, aux situations d'énonciation ou aux contextes social, historique ou politique. Toutefois, un·e linguiste adoptant une approche d'analyse de discours, qui aura pour but d'étudier les textes en relation à ce qui les entoure (cf. Sitri et Barats 2017), devra nécessairement prendre en compte toutes ces informations et les intégrer correctement dans le corpus, sous forme de métadonnées.
- III. Conditions d'**exploitabilité** : les textes du corpus doivent être homogènes (il n'est pas opportun de mélanger différents types de textes sans lien entre eux) et suffisamment

¹⁷ Pour donner un autre exemple, une recherche portant sur les termes géographiques nécessiterait un corpus composé de textes en lien avec la géographie, plutôt que d'articles traitant de mode ou de recettes de cuisine.

nombreux pour permettre l'identification de tendances significatives (plus le corpus est grand, plus le résultats sont solides statistiquement).

Dans le tableau suivant, nous présentons nos corpus de travail pour le français et pour l'italien. Pour chaque langue, nous avons sélectionné un corpus composé d'allocutions et de discours présidentiels, ainsi que deux corpus composés d'articles de presse (un pour la presse nationale et un pour la presse régionale). Les articles pour les corpus de presse sont des versions électroniques, sauf pour le corpus de presse régionale italienne *L'Adige*, composé de versions papier qui ont été océrisées¹⁸.

¹⁸ L'océrisation, également connue sous le terme de « reconnaissance optique de caractères » (de l'anglais *optical character recognition*, OCR), est une technologie qui convertit des documents, comme des images de pages de journal ou de manuscrits anciens, en texte lisible et éditable par les logiciels informatiques. Cette méthode permet d'analyser les caractères présents dans une image et de les transformer en texte numérique, ce qui simplifie la modification, la recherche et l'archivage des informations contenues dans ces documents.

corpus	sigle	types de textes	année(s)	tokens	nombre de textes	source du corpus	
f r a n ç a i s	PRESIDENTS FRANCAIS	PRES-FR	allocutions/discours présidentiels	1958-2020	2843577	733 discours	Hyperbase Web (téléchargé et adapté par Linda Sanvido)
	LE MONDE	LM	presse quotidienne nationale	2010	17895009	37030 articles sélectionnés de façon aléatoire	Europresse (téléchargé et adapté par Cyrielle Montrichard)
	L'EST REPUBLICAIN	ER	presse quotidienne régionale	2010	25825949	tous les articles publiés entre le 2.01.2010 et le 26.02.2010 et entre le 17.03.2010 et le 30.04.2010	Ortolang (téléchargé et adapté par Helena Bermúdez Sabel)
i t a l i e n	PRESIDENTI ITALIANI	PRES-IT	allocutions/discours présidentiels	1955-2021	2639007	2194 discours	www.quirinale.it (constitué via webscraping par Linda Sanvido)
	LA REPUBBLICA	REP	presse quotidienne nationale	2010-2015	26418073	52996 articles sélectionnés de façon aléatoire	www.repubblica.it (constitué via webscraping par Linda Sanvido)
	L'ADIGE	AD	presse quotidienne régionale	1999-2006	20343148	38496 articles sélectionnés de façon aléatoire	Corpus CRIPCO (téléchargé et adapté par Helena Bermúdez Sabel)

Tableau 1 : détails des corpus.

Le corpus des allocutions des présidents français (désormais PRES-FR) a été téléchargé de la version Web d'Hyperbase¹⁹ et adapté pour l'utilisation dans la plateforme d'analyse de

¹⁹ Plateforme développée par le laboratoire UMR 7320 : Bases, Corpus, Langage. Responsables scientifiques : Damon Mayaffre et Laurent Vanni. URL : <http://hyperbase.unice.fr/hyperbase/>

données TXM (Heiden et al. 2010)²⁰. Il contient des allocutions et des discours des présidents français de Charles de Gaulle à Emmanuel Macron, pour un total de presque 3 millions de mots. Le corpus de presse nationale (désormais LM) compte environ 18 millions de mots et contient des articles du journal *Le Monde*, qui ont été téléchargés d'Europresse²¹ en format HTML. Les données extraites ont ensuite été mises en forme à l'aide du logiciel RStudio pour une exploitation dans TXM. Le corpus de presse régionale (désormais ER) a été constitué à partir d'articles tirés du journal de Franche-Comté et de Lorraine *L'Est Républicain* et compte environ 26 millions de mots. Les fichiers TEI ont été récupérés sur ORTOLANG²² et ont ensuite été convertis au format TXT pour être importés dans TXM.

Pour la création des corpus italiens, nous avons veillé à ce que chaque corpus soit comparable avec son correspondant français, en termes de type de texte, de tranche temporelle et de taille. Le corpus des allocutions des présidents italiens (désormais PRES-IT) contient des discours des présidents italiens, de Giovanni Gronchi à Sergio Mattarella, pour un total de 2,6 millions de mots. Le corpus de presse nationale (désormais REP) compte 26 millions de mots et est composé d'articles tirés du journal *La Repubblica* ; celui de presse régionale (désormais AD) d'articles de *L'Adige*, journal de la région du Trentino-Alto Adige (20 millions de mots). Les données pour les deux premiers corpus ont été collectées à l'aide de l'extension Web Scraper pour Google Chrome²³, et ensuite mises en forme pour l'exploitation dans TXM à l'aide des logiciels Notepad++ et RStudio. Le corpus de presse régionale est issu du corpus CRIPCO²⁴, mais ses données ont été adaptées pour une utilisation dans TXM.

²⁰ La plateforme TXM a été initiée par le projet ANR Textométrie et est diffusée gratuitement sous licence open-source. Pour plus d'information, se référer à la présentation du projet : <https://txm.gitpages.huma-num.fr/textometrie/Presentation/>

²¹ <https://nouveau.europresse.com/Search/Advanced>

²² Analyse et traitement informatique de la langue française - UMR 7118 (ATILF), Cognition, Langue, Langages, Ergonomie - UMR 5263 (CLLE) (2020). *Corpus journalistique issu de l'Est Républicain* [Corpus]. ORTOLANG (Open Resources and TOols for LANGuage) - www.ortolang.fr, v4, https://hdl.handle.net/11403/est_republicain/v4

²³ WebScraper fait partie des outils de *Web scraping*, qui permettent d'extraire automatiquement des données des sites internet : <https://webscraper.io/>

²⁴ Corpus CRIPCO (Cross-document Italian People Coreference) : <http://metashare.elda.org/repository/browse/cripco/4ab6b4560f8d11e2b7c50015c5ed7672044ca04410a24fe4b0b894bcaa77075d/>

Tous les corpus français et italiens ont été importés dans TXM au moyen du module `TEXT + CSV`, et ensuite lemmatisés et annotés morphosyntaxiquement de manière automatique avec l'extension TreeTagger²⁵.

3.3 Sélection des paradigmes des verbes d'indication de point de vue

Une fois nos corpus constitués, nous avons adopté une approche *corpus-driven* pour la sélection du paradigme des verbes à étudier (cf. sections 3.3.1 et 3.3.2). En effet, l'adoption d'un paradigme créé *a priori* s'avère problématique car cette façon de procéder présente plusieurs difficultés de délimitation. Si nous partons des listes fournies dans le cadre de la littérature sur les verbes d'indication de point de vue (cf. section 2.2), il faut effectuer un choix. Quelle(s) liste(s) parmi celles proposées faut-il choisir ? Faut-il en croiser plusieurs, de sorte à avoir un inventaire exhaustif ? Mais comment justifier ces choix de façon rigoureuse ?

Une option permettant d'éviter ces problèmes pourrait être de nous baser sur les classifications sémantiques issues de la tradition grammaticale française, du moment que certaines catégories sémantiques de verbes – pensons à ceux qu'on appelle généralement « verbes d'opinion » – se prêtent bien à la fonction d'indicateurs pragmatiques problématisée par Recanati (1984). Cependant, nous constatons que même ces listes sont de taille différente, souvent pas exhaustives, et que les étiquettes employées pour se référer aux familles sémantiques ne sont que partiellement partagées : on parle de verbes d'opinion, de connaissance, de réflexion, d'activité intellectuelle, de verbes psychologiques²⁶. Pour l'italien, ces catégorisations sémantiques sont moins exhaustives, même si elles partagent dans les grands traits les mêmes désignations qu'en français, principalement *verbi d'opinione* (Sabatini 1985, Cortellazzo 1995, Della Valle et Patota 2009, Marelllo 2013) et *verbi psicologici* (Jezek 2003, Savoia 2008, Dal Negro 2015)²⁷.

²⁵ TreeTagger est un outil de traitement automatique de la langue (TAL) qui est conçu pour attribuer de manière automatique une catégorie grammaticale et un lemme à chaque mot présent dans un texte, en se basant sur la langue en question. URL : <https://www.cis.uni-muenchen.de/~schmid/tools/TreeTagger/>

²⁶ Voir, à titre d'exemple, les listes proposées par Willems (1981), Riegel et al. (2002), Abeillé et Godard (2021), ainsi que l'important travail taxinomique de Dubois et Dubois-Charlier (1997).

²⁷ Il manque pour l'italien une classification sémantique systématique comme celle de Dubois et Dubois-Charlier (1997). Les tentatives les plus accomplies nous sont fournies par la linguistique computationnelle et se basent sur les classifications déjà existantes pour l'anglais. Nous mentionnons notamment *ItalWordNet* (Roventini et al.

Ce n'est évidemment pas tâche simple que de définir et nommer des catégories précises, étanches et universelles, vu la multiplication de facteurs qu'il est nécessaire de prendre en compte. Les classifications sémantiques traditionnelles sont basées sur des listes de verbes à l'infinitif et ne tiennent pas compte de tous les sens différents d'un prédicat selon le temps verbal, la personne, ou le contexte d'utilisation. Prenons quelques exemples – tirés de nos corpus – avec le verbe *penser* pour expliquer la diversité d'emplois que ce prédicat peut recouvrir :

- (41) Et quand je dis la France, **je pense à** l'immense majorité des Français.
(PRES-FR)
- (42) Je n'ai atteint la puberté qu'à 16 ans, **je pensais que** j'étais anormal.
(LM)
- (43) Le dissident chinois Liu Xiaobo vient d'être condamné à onze ans de prison pour avoir exercé des droits universellement reconnus : dire et écrire **ce qu'il pense**.(LM)
- (44) « Je m'en fous, **je pense que** tout le monde s'en fout », renchérit un conseiller de l'Elysée. (LM)
- (45) Le public de jazz le fête particulièrement pour sa participation à The Great Guitars, fondé en 1973 avec Charlie Byrd et Barney Kessel, un « supergroupe », où chacun sait ranger son ego pour **penser** la musique collectivement. (LM)

Dans les exemples (41) à (43), le verbe *penser* décrit l'état psychologique du locuteur qui fait référence à son processus mental de réflexion : ce sont des emplois descriptifs au sens de Benveniste (1966). Nous avons au contraire un emploi mitigateur dans (44), car *je pense que* met en scène l'attitude du locuteur vis-à-vis de l'énoncé qu'il profère, en l'occurrence un

2003) et *Italian VerbNet* (Busso et Lenci 2016), qui propose une ébauche de classification des verbes en sept classes sémantiques, même si à l'heure actuelle la catégorie verbes d'opinion/psychologiques n'y figure pas. Plus récemment, Ricchiardi et Ježek (2022) ont effectué une annotation manuelle sur corpus des « propositional attitude verbs » et classent ces prédicats sur la base de leur structure argumentative. Plusieurs groupes sont identifiés, parmi lesquels figurent des verbes de communication (*dire, affermare, ricordare, raccontare, spiegare, ...*), des verbes d'opinion (*pensare, credere*), des verbes d'activité mentale (*verificare, dimostrare, concludere*) et des verbes de perception (*sentire, notare*).

jugement épistémique. Le verbe est superflu au contenu de *p*, il ne fait que caractériser l'attitude du locuteur par rapport à *p*. Rappelons que pour Benveniste (1966), cette manifestation particulière de la subjectivité dans le discours prend forme uniquement à la première personne au repère temporel du présent : l'emploi mitigateur ne sera alors jamais activable avec un pronom autre que *je* – cf. l'exemple (43) – ou un temps autre que le présent de l'indicatif – cf. l'exemple (42). Enfin, l'exemple (45) montre que le verbe *penser* peut également être utilisé comme synonyme de *concevoir*, *travailler*, *jouer*. Ceux que nous venons de présenter ne sont que quelques exemples d'emplois différents d'un même verbe : pour *penser*, le LVF liste dix emplois différents selon le contexte d'utilisation²⁸.

Ces réflexions démontrent ainsi comme le regroupement de verbes en classes sémantiques n'est pas une tâche simple et que les frontières entre les différentes classes ne sont pas forcément toujours très nettes. La question qui se pose est alors la suivante : **comment sélectionner notre paradigme de verbes à étudier sans que celui-ci soit biaisé par une catégorisation sémantique faite à priori ?** Le but de ce travail étant d'étudier les verbes d'indication de point de vue qui sont transparents – au sens de Recanati (1984), autrement dit les verbes qui n'apportent pas de contenu supplémentaire à *p*, mais qui servent uniquement à indiquer comment *p* doit être comprise –, il faut tout d'abord mettre en relief les critères syntaxiques et sémantico-pragmatiques qui font en sorte que ces verbes puissent remplir cette fonction. La propriété de transparence, qui implique qu'un procédé de mitigation est en cours, ne peut être activée que si le locuteur se présente en tant que *je* parlant « ici et maintenant » (cf. Benveniste 1966). Les deux premiers critères de sélection du paradigme des verbes transparents sont donc morphologiques :

- Le verbe doit être utilisé avec un morphème de première personne du singulier, pour qu'il renvoie au locuteur lui-même ;
- Le verbe doit être utilisé au présent de l'indicatif pour qu'il renvoie au moment même de l'énonciation du locuteur.

Du moment que les verbes transparents fonctionnent en tant que préfaces à un contenu *p*, l'autre critère relève de la syntaxe :

²⁸ Le LVF (Les Verbes Français) est un dictionnaire électronique des verbes français, qui registre 25 610 entrées réparties en 14 classes syntactico-sémantiques (cf. Dubois et Dubois-Charlier 1997). Pour accéder aux différents emplois d'un verbe, suivre le lien http://rali.iro.umontreal.ca/LVF_DEM/LVF-DEM-req.html et insérer le verbe en question.

- Le verbe doit être utilisé dans une structure qui introduit une proposition complétive.

Enfin, il faut préciser que pour être transparent, le verbe doit être superflu au niveau du contenu informationnel de l'énoncé qu'il introduit. Autrement dit, son utilisation n'implique pas de changement de valeur de l'énoncé global, car l'information principale est véhiculée par *p*. Le verbe pourrait alors être remplacé par *à mon avis*, sans que le sens global de l'énoncé change drastiquement : ce test de remplacement a donc été effectué afin de départager les verbes d'indication de point de vue. Dans les exemples suivants, seuls les verbes dans (46) et (47) passent le test :

- (46) **J'ai l'impression que** notre réunion d'aujourd'hui revêt une espèce de relief particulier. (PRES-FR)
- (47) **Je pense que** la reprise économique sera très diversifiée selon les continents et les pays. (LM)
- (48) **Je suis heureux que** la majorité qui a désormais la charge de la France comprenne aussi sa jeunesse. (PRES-FR)
- (49) Ce que je sais, c'est que **je me réjouis que** le Secrétaire général ait accepté d'envoyer cette mission. (PRES-FR)
- (50) Je n'ai pas à me plaindre parce que **je sais que** la vie quotidienne des Français est beaucoup plus à plaindre que la mienne. (PRES-FR)

Dans les exemples (46) et (47), les deux verbes *j'ai l'impression que* et *je pense que* modifient la force illocutoire de *p*, en modulant sa valeur de vérité. Il serait possible de remplacer les verbes par *à mon avis* et le sens de l'énoncé global resterait le même : ces deux verbes sont donc transparents. Les exemples (48) à (50) n'ont pas, à nos yeux, les mêmes propriétés. Bien que *je suis heureux que*, *je me réjouis que* et *je sais que* présentent un contenu comme relevant de l'univers subjectif du locuteur, ils ne sont pas transparents au même titre que *j'ai l'impression que* et *je pense que*. En effet, leur rôle n'est pas de modifier la valeur de vérité de *p* : ils ne sont pas superflus au contenu informationnel de l'énoncé, ils y contribuent. En présentant les contenus qu'ils introduisent comme présumés, à savoir comme déjà acquis et partagés dans le *common ground* (cf. Gauker 1998, Stalnaker 2002), ces verbes ne viennent pas modifier la valeur de vérité de *p*, mais deviennent le propos même de l'énoncé : le fait que le locuteur soit heureux que *p* (48), qu'il se réjouisse de *p* (49) et qu'il soit au courant de *p*

(50). Le remplacement du verbe par *à mon avis* entrainerait un changement de sens de l'énoncé global, avec notamment la perte de la présupposition de *p* et la perte de l'information concernant l'état psychologique du locuteur par rapport à *p*.

Les réflexions théoriques qui ont créé les prémisses pour l'établissement d'un paradigme de verbes d'indication de point de vue qui soit *corpus-driven* seront suivies, dans la prochaine section, par la présentation des étapes méthodologiques qui ont été établies pour la création du paradigme.

3.3.1 Extraction du paradigme des verbes français

A partir des critères syntactico-sémantiques qui viennent d'être détaillés, nous nous laissons guider par les corpus PRES-FR, LM et ER, afin de détecter quels sont les verbes transparents qui sont utilisés de façon privilégiée dans ces corpus, en adoptant une approche *corpus-driven*.

Une approche *corpus-driven* vise à dériver de façon systématique des catégories linguistiques à partir de modèles récurrents et de distributions de fréquences qui émergent de la langue dans le contexte (Tognini-Bonelli 2001 : 87), alors qu'une approche *corpus-based* permet avant tout aux linguistes d'utiliser les corpus pour sélectionner des exemples venant exposer, tester ou valider des hypothèses et descriptions théoriques (Tognini-Bonelli 2001 : 65). Dans ce sens, la démarche *corpus-driven* est plus inductive et elle permet de construire des hypothèses non pas à priori, mais à partir des corpus eux-mêmes (Biber 2015 : 196-197). Cette dernière démarche nous semble alors la plus pertinente en ce qui concerne notre question de recherche, puisqu'elle permet de créer une liste de verbes qui ne soit pas soumise au sentiment linguistique personnel du linguiste, mais basée sur des données qui tiennent compte de la fréquence et de la représentation effective des verbes dans les corpus consultés²⁹. Cependant, comme le relèvent De Cesare et al. (2016 : 77), une opposition stricte entre approches *corpus-driven* et *corpus-based* est difficile à maintenir, car même dans les approches les plus radicalement *corpus-driven*, il existe toujours un lien avec les catégories grammaticales traditionnelles. C'est le cas de notre démarche : nous nous laissons guider par les données des corpus afin de construire les paradigmes des verbes d'indication de point de vue, mais nous maintenons le lien avec les catégories grammaticales traditionnelles lorsque nous nous fions à

29

l'étiquetage morphosyntaxique classique (verbes au présents, au participe passé, adjectifs) pour l'extraction des formes.

Notre démarche d'extraction des verbes transparents s'articule en quatre étapes. **La première étape** consiste à isoler les verbes qui sont très fréquemment utilisés avec un pronom personnel de première personne. Pour ce faire, nous avons extrait, dans les trois corpus français, tous les verbes conjugués au présent de l'indicatif qui apparaissent de façon statistiquement significative à côté des morphèmes *je* ou *me*.

Dans TXM, une attraction statistiquement significative entre deux items est indiquée par **l'indice de cooccurrence de Lafon (1981)**. Ces deux items sont plus précisément appelés le **pivot** et le **cooccurrent**. Le pivot est un lemme ou un mot-forme (ou une suite de lemmes ou de mots-formes) à partir duquel la recherche est effectuée (dans notre cas, les verbes d'indication de point de vue). C'est à partir du pivot que l'on obtient ses cooccurrents, à savoir les lemmes ou mots-formes apparaissant de façon statistiquement significative à ses côtés. La recherche des cooccurrents s'effectue en fixant des **empan**s, qui correspondent aux fenêtres d'items qui sont interrogées à la gauche et/ou à la droite du pivot. L'indice de cooccurrence de Lafon (1981) est basé sur un calcul de probabilité et repose sur quatre éléments : la fréquence d'apparition du cooccurrent au sein d'un empan de texte donné, la fréquence totale d'apparition du cooccurrent dans l'ensemble du corpus, la taille de l'empan et la taille du corpus complet. L'indice seuil est fixé à 2 : une rencontre entre un pivot et son cooccurrent marquée par cet indice signifie que le cooccurrent a moins d'une chance sur cent d'apparaître dans l'empan choisi avec une fréquence aussi élevée. En d'autres termes, si la distribution du cooccurrent dans le corpus était aléatoire, une fréquence moins élevée aurait été observée dans 99% des cas. L'indice de cooccurrence permet donc d'identifier une relation statistiquement importante entre deux formes linguistiques. Pour toutes nos recherches, nous avons décidé d'augmenter le seuil de cet indice à 3 et de fixer la cofréquence minimum à 2, afin d'augmenter la représentativité statistique des résultats et d'exclure les hapax³⁰. Ce choix nous permet d'obtenir des indications fiables même pour les verbes qui n'ont pas des fréquences très élevées.

En partant des pivots *je* et *me*, interrogés en tant que lemmes, tous les verbes spécifiques dans un empan droit de deux items ont été extraits. Cet empan est volontairement très réduit pour

³⁰ Les formes dont on ne peut relever qu'une occurrence dans un corpus donné.

ne retenir que les verbes syntaxiquement liés au pivot. Dans la liste de cooccurrents que TXM fournit, seuls les verbes au présent de l'indicatif (VER:pres) ont été gardés. Cette sélection est rendue possible grâce à l'annotation morphosyntaxique effectuée par l'étiqueteur TreeTagger. Voici un aperçu de cette recherche et de ses résultats dans TXM :

Barre de recherche

Pivot: je/me

Paramétrage de l'empan droit (2 items): 2

Paramétrage de la cofréquence minimum: 2

Paramétrage de l'indice de cooccurrence (Lafon 1981): 3,0

Cooccurrent	Fréquence	CoFréquence	Indice	Distance moyenne
VER:pres_rappeler	294	189	196	.6
VER:pres_remercier	206	161	190	.7
VER:impf_dire	578	234	179	.5
VER:pres_faire	3869	549	169	.3
VER:pres_sembler	335	175	159	.0
VER:pres_paraître	315	167	153	.0
VER:pres_aimer	323	168	152	.4
PRO:PER_lui	3639	505	151	.2
VER:ppr_vouloir	664	222	149	1.0
VER:impf_vouloir	412	180	145	.1
VER:futu_faire	875	241	139	.5
VER:impf_évoquer	161	119	135	.1
VER:pres_pouvoir	6426	662	132	.4
VER:pres_tenir	604	196	128	.1
VER:pres_avoir	47298	9775	127	.3
KON_que	46042	4703	127	1.0
ADV_ne	35612	6595	127	.0
PRO:PER_vous	14325	2374	127	.3
PRO:PER_la/le	9886	1853	127	.2
PRO:PER_le	7492	2008	127	.3
PRO:PER_en	6926	1347	127	.2
PRO:PER_me	6899	2061	127	.2
VER:pres_aller	5058	1013	127	.2
VER:inf_dire	4958	918	127	.9
VER:pres_vouloir	4914	1980	127	.2
VER:impf_avoir	4352	916	127	.2

Résultats

Exemples de verbes non retenus (autres modes et temps verbaux): VER:impf_dire, VER:impf_vouloir, VER:futu_faire, VER:impf_évoquer, VER:pres_pouvoir, VER:pres_tenir, VER:pres_avoir, KON_que, ADV_ne, PRO:PER_vous, PRO:PER_la/le, PRO:PER_le, PRO:PER_en, PRO:PER_me, VER:pres_aller, VER:inf_dire, VER:pres_vouloir, VER:impf_avoir

Exemples de verbes retenus (présent de l'indicatif): VER:pres_rappeler, VER:pres_remercier, VER:pres_faire, VER:pres_sembler, VER:pres_paraître, VER:pres_aimer

Figure 5 : exemple de recherche dans TXM.

La Figure 5 présente l'interface de l'outil « cooccurrences » de TXM. L'onglet en haut à gauche correspond au corpus sélectionné (dans ce cas PRES-FR). La barre grise permet de paramétrer la requête en spécifiant le pivot et les propriétés de ses cooccurrents : ici, nous avons interrogé le contexte droit du pivot dans un empan de deux items, spécifié une fréquence minimum d'apparition de chaque cooccurrent dans le corpus (2), une cofréquence minimum

entre le cooccurrent et le pivot (2), et un seuil pour l'indice de spécificité (3). De plus, nous avons choisi de regrouper les cooccurrents en lemmes (`frlemma`) et d'afficher leur catégorie morphosyntaxique (`frpos`). Dans la partie blanche est présentée la liste des résultats : pour chaque cooccurrent sont indiquées sa fréquence d'apparition dans le corpus, sa cofréquence avec le pivot, la valeur de l'indice de spécificité de cette cofréquence et la distance moyenne en termes d'empan entre le cooccurrent et le pivot. La même recherche a été effectuée sur nos trois corpus (PRES-FR, LM, ER) et les résultats ont été croisés, de façon à ne retenir que les verbes communs à l'ensemble des corpus.

La deuxième étape consiste à extraire tous les verbes spécifiques à *que* ou à *si* : ces deux conjonctions nous permettent de repérer les verbes associés à des constructions complétives. En partant des pivots *que* et *si*, en tant que lemmes, nous interrogeons cette fois leur empan gauche, de deux items, et retenons les verbes qui leur sont associés de façon statistiquement significative dans les trois corpus. La démarche est effectuée dans les trois corpus et les résultats croisés comme pour la première étape.

La troisième étape consiste à comparer le tableau des verbes spécifiques à *je/me* et le tableau des verbes spécifiques à *que/si* pour ne retenir que les verbes spécifiques à *je/me* et *que/si* en même temps. L'analyse croisée de ces données dans les trois corpus nous permet ainsi d'identifier les verbes associés de façon prototypique à un morphème de première personne pour introduire une proposition complétive.

Une quatrième et dernière étape s'avère indispensable pour repérer des prédicats complexes tels que *j'ai l'impression que*, *j'ai le sentiment que*, *je suis convaincu-e que*, qui nous semblent également tout autant pertinents pour notre question de recherche, du moment que ces expressions sont également transparentes. L'étiqueteur TreeTagger ne permettant pas de les repérer automatiquement, des recherches supplémentaires ont dû être menées. Nous avons ainsi interrogé les pivots suivants :

```
[frlemma="je"] []? [word="suis"]
```

```
[frlemma="je"] []? [word="ai"]
```

Le []? permet d'accepter un éventuel item entre le pronom personnel et l'auxiliaire (notamment le *ne* de la négation). Pour le premier pivot, nous avons extrait la liste des cooccurrents spécifiques dans un empan droit de deux items, en ne retenant que les adjectifs et les verbes au participe passé et présent. Pour le deuxième pivot, nous avons sélectionné tous les noms spécifiques dans un empan droit de trois items. L'augmentation de l'empan se justifie

par le fait qu'il y aura forcément un déterminant entre l'auxiliaire et le nom, en plus d'un potentiel *pas* de négation. Une fois toutes les expressions repérées, nous avons retenu uniquement celles étant spécifiques à *que/si*. Ces démarches ont été entreprises pour les trois corpus et les résultats croisés à nouveau.

Les quatre étapes méthodologiques permettent ainsi d'obtenir une liste contenant 51 prédicats. Il s'agit donc des prédicats qui, dans les trois corpus, partagent le fait d'être spécifiques en même temps à *je/me* et à *que/si* :

accepte, ajoute, (m')aperçois, attends, ai conscience, ai l'impression, ai la conviction, ai le sentiment, ai peur, avoue, comprends, confirme, considère, constate, crains, crois, (me) demande, dis, doute, entends, espère, m'étonne, suis conscient·e, suis convaincu·e, suis heureux·euse, suis persuadé·e, suis sûr·e, ignore, imagine, note, observe, oublie, pense, précise, (me) rappelle, reconnais, regrette, me réjouis, remarque, répète, sais, semble, sens, souhaite, me souviens, trouve, vois, veux

A partir de cette liste, seuls les verbes transparents sont gardés : pour les départager, le test de *à mon avis* a été effectué (cf. section 3.3). De plus, pour des raisons de représentation statistique, nous avons retenu uniquement les séquences apparaissant avec une fréquence brute minimum de 25 occurrences dans au moins PRES-FR et LM³¹. Les tableaux indiquant les verbes exclus à cause de l'un ou de l'autre de ces deux critères sont présentés dans le Tableau 7 des Annexes.

Dans le Tableau 2 ci-dessous, nous présentons le paradigme final des verbes d'indication de point de vue retenus pour le français. Chaque forme est accompagnée par la recherche *Corpus Query Language* (CQL), que nous avons utilisée pour calculer sa fréquence et pour mener les analyses de ses cooccurrents (cf. section 3.3.3). Dans la partie droite du tableau, se trouvent les fréquences brutes (Fb) et les fréquences relatives (Fr) de la forme selon le corpus dans lequel elle intervient. Les fréquences brutes, dites aussi fréquences absolues, correspondent au nombre total d'occurrences de la forme dans le corpus ; les fréquences

³¹ Le choix de trier les séquences *je/il me + syntagme verbal + que* par rapport à leurs fréquences de façon distincte par rapport à chaque corpus, n'aurait pas permis d'avoir un paradigme à la fois riche et identique d'un corpus à l'autre : en effet, les seuls verbes qui ont une fréquence brute d'au moins 25 items dans les trois corpus sont *je crois que*, *je pense que* et *je trouve que*. Le choix a donc été pris, pour les deux langues, de départager les verbes sur la base de leurs fréquences dans deux corpus uniquement (les corpus des allocutions présidentielles et ceux de presse nationale) ; l'analyse cooccurrentielle (cf. chapitre 4) est d'ailleurs menée uniquement sur ces deux discours.

relatives indiquent le rapport³² entre la fréquence brute et le nombre total de mots du corpus, ce qui permet de comparer la fréquence de la forme entre corpus de taille différente (cf. section 3.3.4).

Verbes français	Paramétrage de la recherche dans TXM	PRES-FR		ER		LM	
		Fb	Fr	Fb	Fr	Fb	Fr
je crois que	[frlemma="je"] [word="crois"] [frlemma="que"]	1056	371	195	8	250	14
je pense que	[frlemma="je"] [word="pense"] [frlemma="que"]	878	309	558	22	351	20
je considère que	[frlemma="je"] [word="considère"] [frlemma="que"]	202	71	18	1	30	2
je dis que	[frlemma="je"] [word="dis"] [frlemma="que"]	102	36	14	1	26	1
je suis sûr·e que	[frlemma="je"] [word="suis"] [word="sur sûr sûre"] [frlemma="que"]	90	32	15	1	29	2
je trouve que	[frlemma="je"] [word="trouve"] [frlemma="que"]	90	32	62	2	38	2
je suis convaincu·e que	[frlemma="je"] [word="suis"] [word="convaincu convaincue"] [frlemma="que"]	68	24	11	0	33	2
il me semble que	[word="il"%c] [word="me"] [word="semble"] [frlemma="que"]	55	19	17	1	42	2

Tableau 2 : paradigme des verbes d'indication de point de vue français.

La méthodologie *corpus-driven* que nous adoptons pour cette étude a l'avantage de construire un paradigme en s'appuyant sur des données réelles, issues de corpus différents. Les verbes

³² Nous avons décidé d'exprimer ce rapport en nombre d'occurrences par million de mots, mais il peut également être exprimé en pourcentage.

d'indication de point de vue ne sont donc pas sélectionnés en partant de catégorisations subjectives et questionnables, mais sur la base de leur représentativité en termes de fréquences dans des contextes authentiques. Ce paradigme présente ainsi des verbes dont l'usage est systématisé dans des structures syntactico-sémantiques très précises : ce sont des verbes utilisés de manière prototypique au présent de l'indicatif, avec un morphème de première personne, introducteurs de complétive sur laquelle ils effectuent un commentaire qui n'apporte aucun contenu informationnel supplémentaire à l'énoncé. En d'autres mots, ce sont des verbes qui interviennent pour modaliser la complétive qu'ils introduisent.

3.3.2 Extraction du paradigme des verbes italiens

Pour la sélection des verbes italiens, nous avons utilisé la même méthodologie que pour le français. Le choix de ne pas construire le paradigme en partant des traductions des verbes français s'explique par la volonté d'éviter une démarche qui pourrait se voir influencée et biaisée par les résultats propres aux verbes et aux corpus français. Quelques ajustements ont néanmoins dû être effectués à cause de la morphosyntaxe de l'italien. Nous n'avons notamment pas pu partir d'une recherche automatique de tous les verbes spécifiques au morphème de première personne *io* (« je ») car, comme ce pronom personnel n'est pas obligatoire en italien, cette recherche n'aurait pas permis l'inclusion de tous les verbes utilisés sans pronom explicite³³.

Ce sont donc les conjonctions *che/se* (« que/si ») qui ont été fixées comme pivot de départ. A partir de ces pivots, tous les cooccurents dans un empan gauche d'un item ont été extraits (l'empan ne nécessite pas d'être élargi car en italien il n'y a aucun élément de négation entre le verbe et la conjonction). Une fois chaque cooccurent affiché en `frpos` et en `word`, il a été possible de sélectionner manuellement tous les items conjugués à la première personne du singulier – ou à la troisième du singulier pour les verbes impersonnels tels que *sembrare* – des catégories morphosyntaxiques suivantes :

- `VER:pres` → cette étiquette permet de sélectionner tous les verbes conjugués au présent de l'indicatif

³³ Par exemple, *sono sicuro/a che* n'aurait pas pu être repéré car seulement 3/60 occurrences de *sono sicuro/a che* sont précédées par *io* dans PRES-IT, 2/111 dans AD et 7/307 dans REP.

- ADJ → cette étiquette permet de sélectionner tous les adjectifs qui pourraient faire partie d'expressions telles que *sono sicuro che*
- VER:pper → cette étiquette permet de sélectionner les participes passés qui pourraient faire partie d'expressions telles que *sono convinto che, sono persuaso che*
- NOM → cette étiquette permet de sélectionner des noms faisant également partie d'expressions indiquant une attitude, comme *ho l'impressione che*.

L'affichage en word permet de mettre en relief les marques morphologiques des verbes et des adjectifs, mais il débouche sur des listes plus longues, comme illustré dans l'image suivante :

Requête [itemma="che"]

Paramètres

Propriétés des cooccurrents : itpos ; word Éditer Seuils : Fmin ≥ 2 Cmin ≥ 2 Indice ≥ 3,0

Contexte : forme structure s Utiliser le contexte gauche Utiliser le contexte droit inclure la structure contenant le pivot dans le décompte

de - 0 à - 0 et de 0 à 0

Cooccurrent	Fréquence	CoFréquence	Indice	Distance moyenne
VER:pres_pare	1357	274	219	.0
VER:pres_Pare	144	86	118	.0
VER:pres_pensa	2051	302	200	.0
VER:pres_Pensa	83	34	40	.0
VER:pres_pensano	466	70	47	.0
VER:pres_Pensano	6	3	4	.0
VER:pres_pensate	54	13	12	.0
VER:pres_Pensate	27	17	24	.0
VER:pres_pensi	163	29	22	.0
VER:pres_Pensi	36	18	23	.0
VER:pres_pensiamo	395	88	75	.0
VER:pres_Pensiamo	125	27	23	.0
VER:pres_penso	1756	667	127	.0
VER:pres_Penso	770	421	127	.0
VER:pres_Pereyra	16	5	5	.0
VER:pres_pilota	53	12	11	.0
VER:pres_posso	1725	86	22	.0
VER:pres_precisa	821	148	111	.0
VER:pres_precisano	97	18	14	.0
VER:pres_precisiamo	6	4	6	.0
VER:pres_predichiamo	2	2	3	.0
VER:pres_preferisco	193	11	3	.0
VER:pres_Premiato	44	7	10	.0

t pivot 369131, v cooc 3380, t cooc 256933, T corpus 26418073

Figure 6 : exemple d'affichage des cooccurrents en word.

Avec cet affichage, toutes les personnes du présent de l'indicatif sont visibles, ce qui implique que pour sélectionner uniquement les formes à la première personne du singulier, un tri minutieux est nécessaire. A titre d'exemple, dans les onze occurrences mises en relief dans la Figure 6, seules les deux dernières ont été retenues.

Comme pour le français, les recherches ont été effectuées sur les trois corpus italiens (PRES-IT, AD, REP) et les résultats croisés. La liste suivante présente tous les items qui sont en association spécifique avec la conjonction *che* (aucun élément en association avec *se* n'a été identifié) :

mi accorgo, aggiungo, assicuro, (mi) auguro, auspico, augurio, auspicio, certezza, consapevolezza, convincimento, convinzione, fiducia, impressione, percezione, preoccupazione, sensazione, sentimento, speranza, timore, capisco, confido, credo, dico, dubito, consapevole, contento/a, convinto/a, fiducioso/a, sicuro/a, immagino, osservo, pare, penso, (mi) ricordo, rispondo, ritengo, sembra, sento, so, spero, suppongo, temo, trovo, vedo

Toutefois, dans cette liste, les items suivants sont des éléments qui ne sont pas pertinents dans le cadre de l'extraction des verbes transparents :

augurio, auspicio, certezza, consapevolezza, convincimento, convinzione, percezione, preoccupazione, sensazione, sentimento, timore

Le choix de ne pas inclure le pronom personnel dans la recherche entraîne nécessairement la présence de « bruit », ce qui mène à des imprécisions et doit être corrigé. Observons ces exemples :

(51) La **percezione che** Europa e Islam esprimano incompatibilità è errata. (PRES-IT)

La perception selon laquelle l'Europe et l'Islam sont incompatibles est erronée.³⁴

(52) In qualche modo so che possono essere dei messaggi nella bottiglia, senza **certezza che** vengano raccolti. (PRES-IT)

D'une certaine manière, je sais qu'il peut s'agir de messages en bouteille, sans certitude qu'ils seront ramassés.

Dans (51) et (52), les noms *percezione* et *certezza* ne sont pas utilisés, comme nous l'aurions attendu, dans des expressions telles que *ho la percezione/certezza che*. Tous les éléments que nous avons listés comme non pertinents interviennent, dans les corpus interrogés, en tant que

³⁴ Toutes les traductions italien>français ont été produites de façon automatique au moyen de DeepL, puis adaptées par l'auteurice.

syntagmes prépositionnels : ils doivent donc être exclus. Les recherches pour l'italien impliquent certes un peu plus de tri et de nettoyage manuel que celles pour le français, mais ces étapes sont nécessaires pour la systématique de la démarche qui évite ainsi de passer à côté de certaines expressions qui pourraient se révéler pertinentes.

Une fois cette étape de nettoyage terminée, seuls les verbes qui passent le test de la transparence sont gardés. Dans les deux extraits suivants, le verbe est transparent dans (53), mais pas dans (54) :

- (53) La discussione tra di noi è stata sommamente fruttuosa e io **penso che** lo sarebbe stata anche se avessimo potuto per altre due ore entrare nel merito di molte altre questioni su cui la sintonia tra i due Paesi è molto forte e si rafforzerà ulteriormente. (PRES-IT)

La discussion entre nous a été extrêmement fructueuse et je pense qu'elle l'aurait été même si nous avions pu passer deux heures de plus à discuter de nombreuses autres questions, sur lesquelles l'harmonie entre les deux pays est très forte et sera encore renforcée.

- (54) Eileen è probabilmente la più antica cliente del Paese, e **spero che** anche la sua famiglia proseguirà questa tradizione. (REP)

Eileen est probablement la plus ancienne cliente du pays, et j'espère que sa famille aussi perpétuera cette tradition.

Pour départager les verbes transparents et les verbes non transparents, nous avons effectué le test de remplacement par *secondo me*. Dans (53) le test aboutit : la suite *secondo me lo sarebbe stata anche se...* peut venir se substituer à *penso che lo sarebbe stata anche se...* sans que le sens de l'énoncé global change : le but du locuteur est dans les deux cas d'avancer une supposition concernant les bénéfices d'une conversation de longue durée. Dans (54), le remplacement de *spero che* par *secondo me* n'est pas heureux. En effet, ce qui est exprimé dans *p* change selon le marqueur qui est utilisé : si *spero che* met en lumière le souhait du locuteur concernant les habitudes de la famille d'Eileen (en l'occurrence, d'acheter des voitures de la marque Ford), *secondo me* efface cette mise en scène du désir du locuteur et l'interprétation devient une supposition sur le fait que la famille pourrait, dans le futur, poursuivre cette tradition.

La dernière étape de sélection concerne, comme pour le français, les fréquences d'apparition des formes verbales. Seuls les verbes transparents ayant une fréquence brute minimum de 25

dans le corpus des discours présidentiels (PRES-IT) et dans le corpus de presse nationale (REP) ont été retenus (plus de détails dans le Tableau 8 des Annexes). **Le Tableau 3 ci-dessous présente le paradigme final des verbes d'indication de point de vue retenus pour l'italien, en les accompagnant des fréquences brutes et relatives dans les trois corpus :**

Verbes italiens	Paramétrage de la recherche dans TXM	PRES-IT		AD		REP	
		Fb	Fr	Fb	Fr	Fb	Fr
		credo che	[word="credo"%c] [itlemma="che"]	553	210	2312	114
penso che	[word="penso"%c] [itlemma="che"]	223	85	954	47	1134	43
sono convinto/a che	[word="sono"%c] [word="convinto convinta"] [itlemma="che"]	164	62	367	18	403	15
ritengo che	[word="ritengo"%c] [itlemma="che"]	97	37	318	16	171	6
mi pare che	[word="mi"%c] [word="pare"] [itlemma="che"]	74	28	345	17	114	4
sono sicuro/a che	[word="sono"%c] [word="sicuro sicura"] [itlemma="che"]	60	23	111	5	307	12
dico che	[word="dico"%c] [itlemma="che"]	26	10	309	15	286	11

Tableau 3 : paradigme des verbes d'indication de point de vue italiens.

Nous remarquons que nous retrouvons en tête de liste, avec les fréquences les plus élevées du paradigme, les mêmes verbes que nous avons en français, à savoir *credo che* et *penso che* (*je crois que* et *je pense que*). La majorité des autres verbes recensés sont également les correspondants morphologiques des formes françaises : *sono convinto/a che* (*je suis convaincu·e que*), *sono sicuro/a che* (*je suis sûr·e que*), *dico che* (*je dis que*). La forme impersonnelle *mi pare che* est également présente, même si ce n'est pas la même qu'en français (*il me semble que*). En revanche, le correspondant morphologique de *je considère que*

(*considero che*) n'apparaît pas dans le paradigme des verbes italiens ; toutefois, nous relevons la présence de *ritengo che*, traduction la plus courante de *je considère que*³⁵. Enfin, l'absence de *trovo che* (*je trouve que*) s'explique par la faible fréquence de cette expression dans les corpus pris en compte (cf. Tableau 8 dans les Annexes).

Le fait qu'on arrive à cette homogénéité entre français et italien est un résultat assez frappant, du moment que la démarche utilisée pour extraire les verbes d'indication de point de vue est complètement indépendante et menée sur des corpus différents. Ceci apporte un éclairage sur la façon dont le locuteur se met en jeu pour exprimer son point de vue : les verbes utilisés à ces fins convergent en français et en italien. D'un point de vue purement sémantique, la majorité des verbes appartiennent à la catégorie des verbes d'opinion, à l'exception d'un verbe de parole (*dire/dire*). Si l'on se réfère aux classifications fondées sur la parenthéticité de ces prédicats, on observe que presque tous peuvent fonctionner comme recteurs faibles, à l'exception de *considérer/ritenere* et de *dire* pour le français, pour lesquels l'emploi en incise serait plus difficilement acceptable.

L'analyse des Tableau 2 et Tableau 3 suggère qu'un éventuel choix de construire le paradigme des verbes italiens sur la base des traductions des verbes français aurait été problématique pour au moins deux raisons. Tout d'abord, les correspondants morphologiques de *je considère que* et *je trouve que* ne sont pas très fréquents dans nos corpus italiens : *considero che* apparaît 2 fois dans PRES-IT, 1 fois dans AD et 3 fois dans REP ; *trovo che* respectivement 11, 39 et 21 fois. Ceci montre que nous ne pouvons pas partir du présupposé qu'un correspondant morphologique aura forcément le même usage et la même représentation dans deux langues de même souche. Une deuxième raison est que certains verbes n'auraient pas été retenus, puisqu'ils n'ont pas de correspondants morphologiques dans le paradigme français. Nous nous référons notamment à *ritengo che* (qui est le quatrième verbe le plus fréquent du paradigme italien), et à *mi pare che* (rappelons que pour le français, nous avons uniquement *il me semble*

³⁵ Pour déterminer la traduction la plus fréquente de *je considère que*, nous avons consulté les dictionnaires bilingues en ligne Larousse (<https://www.larousse.fr/dictionnaires/francais-italien/consid%C3%A9rer/15863>) et il Sansoni Francese (https://dizionari.corriere.it/dizionario_francese/Francese/C/considerer.shtml), ainsi que les corpus parallèles OPUS (<https://www.sketchengine.eu/opus-parallel-corpora/>) et Reverso (<https://context.reverso.net/traduzione/francese-italiano/je+consid%C3%A8re+que>).

que et non *il me parait que*). Le même type de problème aurait pu se poser si nous étions parties d'un paradigme de verbes italiens pour le transposer en français.

En somme, la méthodologie *corpus-driven* appliquée de la même façon pour le français et l'italien a permis de mettre en relief à la fois des convergences et des divergences dans la représentativité des verbes d'indication de point de vue, lesquelles auraient échappé à une approche fondée uniquement sur les correspondances morphologiques entre les deux langues.

3.3.3 Optimisation du repérage des verbes dans nos corpus

Dans cette section, nous expliquons les choix méthodologiques qui ont permis, une fois les paradigmes établis, une optimisation des démarches de repérage de chaque forme verbale dans nos corpus.

La colonne « Paramétrage de la recherche dans TXM » des Tableau 2 et Tableau 3 indique que cette extraction a été effectuée au moyen du langage d'expression de requêtes CQL, acronyme pour *Corpus Query Language*. Il s'agit d'un code partagé, utilisé pour rechercher des séquences linguistiques plus ou moins complexes. Une expression CQL comprend une chaîne de caractères qui peuvent correspondre à un ou à plusieurs mots, définis en fonction de leur forme graphique, de leur forme lemmatisée ou de leur catégorie morphosyntaxique. Les requêtes peuvent être combinées à l'aide de ce qu'on appelle des opérateurs booléens, à savoir des symboles comme « & », « | », « ! » ou « ? », qui permettent de filtrer les résultats en combinant ou en excluant des chaînes de caractères. Des modificateurs peuvent également être utilisés, comme « %c » ou « %d » : le premier rend la requête insensible à la casse (majuscules, minuscules), le deuxième aux diacritiques (accents, cédilles, etc.). Ainsi, avec une requête telle que `[word="penser|croire"%c]`, six formes peuvent être repérées de façon simultanée : *penser*, *croire*, *Penser*, *Croire*, *PENSER*, *CROIRE*. Si à la place de l'indicateur `word` nous utilisions `frlemma`, les résultats seraient encore plus nombreux, car nous aurions accès à chaque mot-forme des deux lemmes, à savoir à toutes les conjugaisons et déclinaisons possibles de ces deux verbes.

Toutes les recherches qui seront présentées dans la suite de ce travail ont été effectuées en partant des pivots explicités dans les Tableau 2 et Tableau 3. Ce choix méthodologique ne permet pas le repérage de séquences plus larges, dans lesquelles d'autres formes linguistiques interviendraient entre le verbe et la conjonction, comme dans les exemples suivants :

- (55) **Je pense simplement que** le Premier ministre a raison d'adopter ce nouveau rythme. (PRES-FR)
- (56) Mais **je pense, comme vous, qu'**il faut ajouter toujours à la politique économique des perspectives sociales et, notamment, d'adaptation progressive de notre société. (PRES-FR)

L'adverbe *simplement* et l'incise *comme vous* apportent des précisions par rapport à l'interprétation de la complétive que le verbe *penser* introduit. Le fait d'élargir l'empan entre nos verbes et la conjonction pourrait ainsi nous amener à avoir une variété plus large de formes à analyser, et à voir les différences d'effet rhétorique parmi celles-ci. Toutefois, cette démarche s'avère problématique car l'élargissement du pivot n'assurerait pas la pertinence de toutes les formes. Observons les exemples suivants :

- (57) En même temps, **je crois** beaucoup à ce **que** je fais. (PRES-FR)
- (58) **Je pense** à ces formulaires **que** nous sommes tous obligés de remplir [...]. (PRES-FR)
- (59) J'ai dit, **je crois**, ce **qu'**il fallait dire sur ce sujet. (PRES-FR)

Dans tous ces cas, nous remarquons que le *que* ne travaille plus en accord avec le verbe pour accompagner une proposition complétive. Dans (57) et (58), les verbes *croire* et *penser* sont utilisés avec la préposition *à* pour introduire des groupes nominaux, et le *que* est un pronom relatif reprenant ces GN ; dans (59), le *je crois* est utilisé en emploi parenthétique. Des questionnements se posent également en ce qui concerne des variations syntaxiques d'usage de nos verbes, comme dans les exemples suivants :

- (60) C'est pourquoi **je trouve qu'**à l'heure actuelle il est mauvais de critiquer le comportement et les décisions de ceux qui appliquent l'Etat de droit. (PRES-FR)
- (61) **Je trouve** insensé et immoral **que** la spéculation, procédant à coups de milliards de dollars, puisse obtenir raison contre les Etats qui, eux, représentent les intérêts de leurs peuples [...]. (PRES-FR)

Dans ces deux exemples, le verbe exprime un jugement qui s'explicite par des adjectifs axiologiques tels que *insensé*, *immoral*, *mauvais* (cf. Kerbrat-Orecchioni 1980 : 90). Si dans

(60) *trouver* introduit directement la complétive et est donc comptabilisé dans notre paradigme, dans (61) il introduit les adjectifs *insensé* et *immoral*. Si nous partons du principe que (61) est caractérisé par un phénomène d'ellipse de l'auxiliaire (*Je trouve que c'est insensé et immoral que...*), nous pouvons alors assumer que le verbe *trouver* dans (61) accomplit la même fonction que le verbe *trouver* dans (60). Toutefois, pour des questions de systématisme, nous préférons exclure tous les emplois du type (61).

Afin de maintenir la pertinence et l'uniformité des formes analysées, et pour éviter d'inclure dans le paradigme des constructions non adéquates comme celles relevées dans les exemples (57) à (59), nous avons pris la décision de ne pas élargir le pivot, en ne prenant comme objet d'analyse les séquences où la conjonction *que/che* suit directement le syntagme verbal.

3.3.4 Quelques réflexions sur les fréquences et la représentation des deux paradigmes dans les corpus

Dans les Tableau 2 et Tableau 3, nous avons exposé non seulement les fréquences brutes des verbes d'indication de point de vue, mais aussi les fréquences relatives. Ces dernières permettent de mettre en perspective la place de chaque verbe dans des corpus de taille différente, car le calcul tient compte des occurrences des verbes par million de mots. La fréquence relative est calculée en multipliant par un million le rapport entre la fréquence brute et le nombre total des mots du corpus :

$$F_{rel} = (F_{br} / N_{tot}) * 1000000$$

A titre d'exemple, *credo che* apparaît plus souvent en termes de fréquence brute dans les deux corpus de presse italiens (AD : 2312 et REP : 2114) que dans le corpus des discours présidentiels (PRES-IT : 553). Toutefois, si l'on ramène la fréquence de *credo che* à la taille du corpus, on observe que pour 1 million d'occurrences, la fréquence est de 210 dans PRES-IT, de 114 dans AD et de 80 dans REP. Le calcul des fréquences relatives permet donc de comparer la fréquence d'une forme donnée à travers des corpus de taille différente. Cependant, les fréquences relatives ne fournissent pas d'indication sur la probabilité qu'une forme donnée intervienne autant ou aussi peu de fois dans un corpus par rapport à un autre. En d'autres termes, il n'est pas possible de déterminer si c'est « banal » d'obtenir une telle distribution. Comme l'explique Mayaffre (1997) :

L'emploi d'un mot peut être qualifié d'important ou de faible uniquement par rapport à une norme. Cette norme est une norme endogène, elle est constituée par le corpus lui-même dans son ensemble. Ainsi si le corpus est segmenté [...] on pourra comparer l'utilisation du mot dans une sous-partie par rapport à l'utilisation "moyenne" ou générale dans le corpus. Il s'en suit une série de calculs qui aboutissent à l'établissement d'un écart réduit et qui vise à mesurer le degré de "*spécificité*" d'un terme dans une sous-partie. (Mayaffre 1997 : 170)

Afin d'interroger la ventilation des verbes d'indication de point de vue de notre paradigme, nous souhaitons dégager, dans les deux langues, leurs écarts d'emploi dans les trois corpus pris en compte, afin de vérifier si certains verbes y sont sur ou sous-employés. Pour que cela soit possible dans TXM, il a été nécessaire de réunir les trois corpus du français dans un macro-corpus, puis de partitionner ce macro-corpus en sous-parties. Chaque sous-partie correspond aux corpus PRES-FR, ER, LM. Le même procédé a été entrepris pour les corpus italiens.

Une fois les deux macro-corpus partitionnés, nous utilisons l'**indice de spécificité de Lafon (1980)** pour vérifier la distribution des verbes d'indication de point de vue dans les différentes sous-parties. Le calcul de l'indice de spécificité prend en compte la fréquence globale de la forme dans le macro-corpus, la fréquence de la forme dans chaque sous-partie, la taille globale du macro-corpus et la taille des sous-parties. De plus, il intègre « la part de hasard dans la distribution et mesure la probabilité que telle ou telle fréquence apparaisse dans telle ou telle partie : l'écart par rapport à un tirage aléatoire » (Mayaffre 1997 : 171). En d'autres termes, nous pouvons mesurer la valeur des écarts entre le « poids » d'une forme donnée dans les différentes sous-parties.

Nous présentons dans les Figure 7 et Figure 8 ci-dessous les résultats des recherches menées au moyen de l'indice de spécificité de Lafon (1980) et les commentons par la suite.

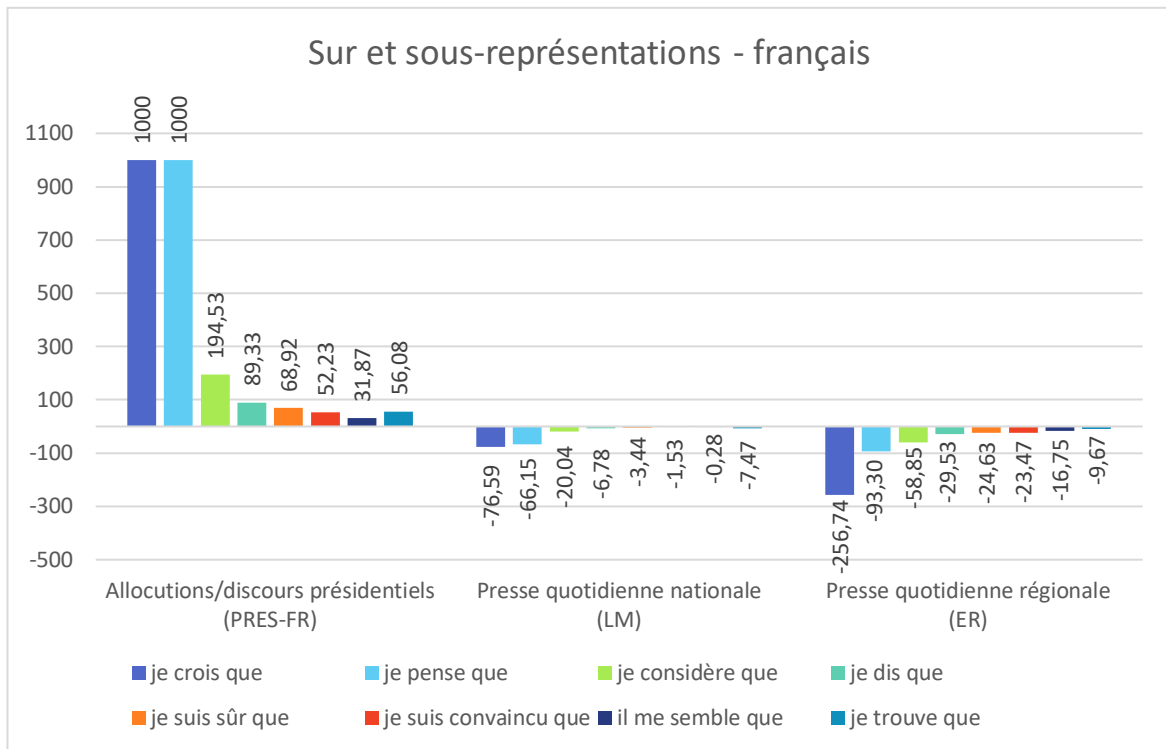


Figure 7 : sur et sous-représentations des verbes d'indication de point de vue dans les corpus français.

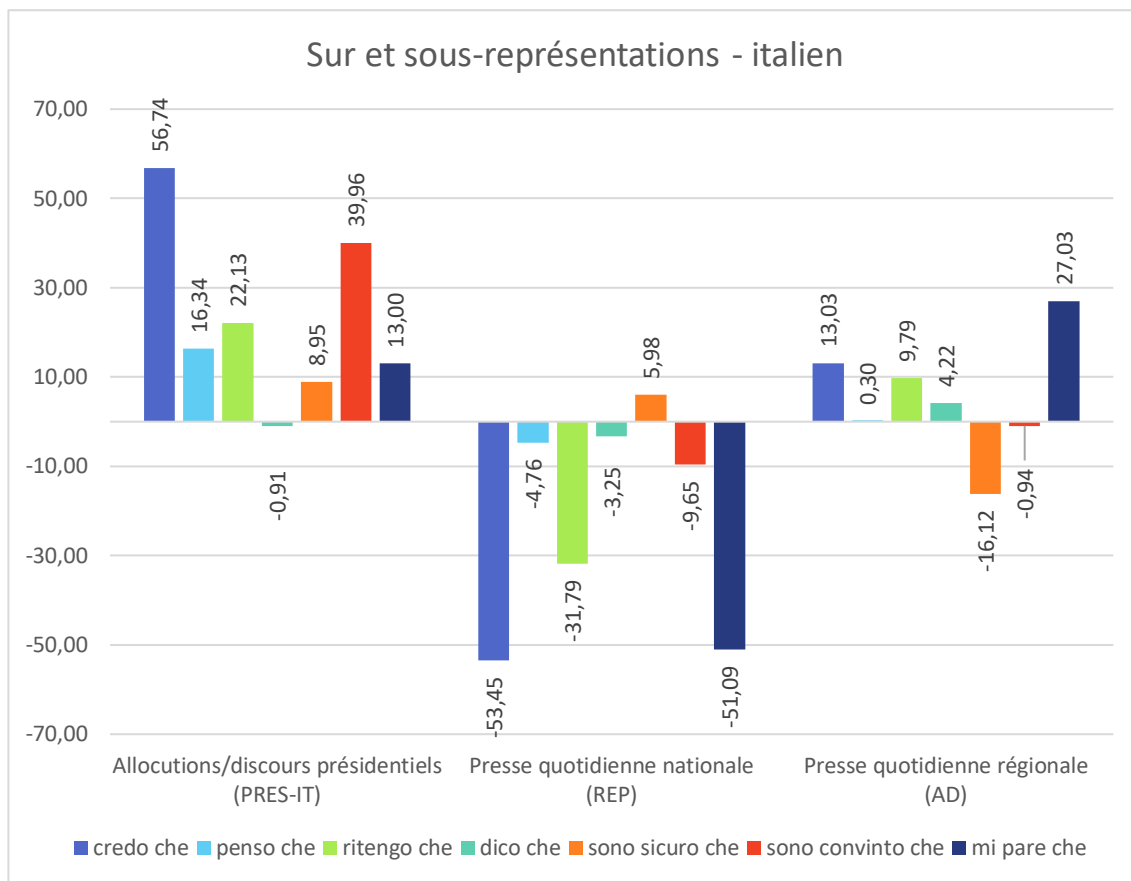


Figure 8 : sur et sous-représentations des verbes d'indication de point de vue dans les corpus italiens.

L'interprétation des tableaux se fait de la manière suivante : tout verbe d'indication de point de vue ayant un indice égal ou supérieur à 3 est considéré comme surreprésenté³⁶. L'indice indique la probabilité que la distribution telle qu'observée soit forte ou faible. Ainsi, un indice de 3 indique qu'il y avait 1 chance sur 1000 pour que la distribution de la forme se présente autant de fois dans la partie. Cet indice révèle donc que le fait que cette forme soit présente autant de fois ne peut être dû au hasard et est donc à imputer à un autre élément (type de texte, fonctionnement en langue, etc.). A l'inverse, un indice égal ou inférieur à -3 indique que le verbe d'indication de point de vue est sous-employé : il apparaît, de manière inattendue au vu de sa répartition probabiliste normale dans le macro-corpus, moins fréquemment que prévu dans la partie. Enfin, lorsque l'indice de spécificité est compris entre 3 et -3, le verbe est

³⁶ L'indice seuil préétabli dans TXM correspond à 2. Nous l'avons porté à 3 afin d'obtenir une représentation statistique plus solide et de nous aligner aux autres recherches menées dans le cadre de cette thèse (cf. section 3.3.1).

considéré comme ni surreprésenté ni sous-représenté (on dit aussi qu'il est en situation de banalité).

La Figure 7 met en évidence deux points principaux pour ce qui concerne les verbes français :

- Lorsque les verbes d'indication de point de vue sont surreprésentés, ils le sont tous et lorsqu'ils sont sous-représentés, ils le sont également tous, ou presque. Ces tendances sont très nettes pour chacun des corpus étudiés.
- Les verbes d'indication de point de vue sont systématiquement surreprésentés dans le corpus des discours présidentiels et sous-représentés dans les deux corpus de presse (à l'exception de *je suis convaincu-e que* et *il me semble que* dans LM, qui sont en situation de banalité).

Ces données démontrent que ces verbes constituent une catégorie homogène, car ils ont tous le même type de profil statistique selon le corpus interrogé. Cependant, il n'existe aucune raison apparente pour que ces formes se comportent de manière uniforme et qu'il y ait une convergence dans leur sur ou sous-représentation selon le corpus. En effet, on pourrait imaginer que certains verbes, tels que *je considère que*, tendraient à caractériser davantage des textes argumentatifs et d'un registre plus formel, comme ceux qui composent le corpus des discours présidentiels. Les résultats montrent pourtant de manière très claire que ces verbes partagent des caractéristiques communes dans leur fonctionnement.

La Figure 8, qui présente les résultats pour les corpus italiens, révèle des tendances similaires à celles observées pour le français, notamment en ce qui concerne le corpus des discours présidentiels et le corpus de presse nationale. Dans PRES-IT, tous les verbes d'indication de point de vue sont surreprésentés, à l'exception de *dico che*, qui est en situation de banalité. Dans REP, tous les verbes sont sous-représentés, à l'exception de *sono sicuro/a che*, qui est surreprésenté. Ces données indiquent que, en italien également, ces formes ont tendance à former une famille qui se comporte de façon cohérente selon le corpus dans lequel elles sont mobilisées, à l'exception toutefois de deux verbes (*dico che* et *sono sicuro/a che*) qui semblent se distancier de cette classe. L'analyse cooccurentielle menée dans le chapitre 5 permettra de mieux comprendre ces résultats. En ce qui concerne le corpus de presse régionale (AD), nous ne pouvons pas généraliser le comportement de ce groupe de verbes car les résultats sont beaucoup moins homogènes, avec quatre verbes surreprésentés, un sous-représenté et deux en

situation de banalité. L'établissement du profil combinatoire pour ces formes dans AD pourrait aider à expliquer ces données³⁷.

En conclusion, l'analyse statistique des sur et sous-représentations des verbes appartenant aux deux paradigmes a mis en évidence que ces formes constituent un ensemble cohérent. En effet, elles fonctionnent de manière similaire selon le corpus dans lequel elles sont mobilisées, ce qui justifie leur analyse conjointe en tant que membres d'une classe plus large de « verbes d'indication de point de vue ». De plus, elle a démontré la nécessité de ne pas se fier exclusivement aux fréquences relatives, en offrant un nouvel angle de lecture avec un outil statistique fondé sur la probabilité et permettant de confronter les formes dans chacun des corpus, les unes par rapport aux autres. Reprenons l'exemple de *credo che* mentionné au début de cette section. Cette forme apparaît comme surreprésentée dans PRES-IT (indice de 56,74) et dans AD (indice de 13,03) mais pas dans REP, où elle est au contraire fortement sous-représentée (indice de -53,45). L'indice de spécificité de Lafon (1980) fournit une clé d'interprétation que la seule fréquence relative ne permet pas d'appréhender pleinement. En effet, si cette dernière indique simplement que *credo che* est plus fréquent dans PRES-IT (fréquence relative de 210) que dans AD (114) et REP (80), l'analyse des indices de spécificité révèle que non seulement *credo che* est moins fréquent dans REP, mais qu'il y est nettement sous-employé par rapport aux deux autres corpus. Un autre exemple concerne *dico che*, qui présente des fréquences relatives très proches dans les trois corpus : 10 dans PRES-IT, 15 dans AD et 11 dans REP. Malgré cela, l'indice de spécificité permet de montrer qu'il est en fait surreprésenté dans AD, sous-représenté dans REP et en situation de banalité dans PRES-IT.

3.4 L'analyse cooccurentielle des verbes d'indication de point de vue

Les recherches qui ont utilisé une approche sur corpus pour étudier les verbes de notre paradigme se sont principalement basées sur les corpus en tant que réservoirs d'exemples, afin de décrire les usages ou les fréquences des verbes. A notre connaissance, aucune étude n'a cherché à décrire ces verbes sur la base de leurs préférences sémantiques, à savoir en tenant compte des mots qui interviennent de façon statistiquement significative dans le voisinage

³⁷ A l'heure actuelle, l'analyse cooccurentielle des verbes d'indication de point de vue français et italiens (sections 4 et 5) a été menée uniquement dans les corpus des discours présidentiels (PRES-FR et PRES-IT) et dans les corpus de presse nationale (LM et REP).

sémantico-syntaxique des verbes d'indication de point de vue et qui contribuent à la construction de leur sens. Une recherche s'intéressant aux régularités dans l'exploitation de ces verbes en discours, à travers l'étude des formes qui leur sont associées de façon statistiquement significative (leurs cooccurents spécifiques), nous semble dès lors nécessaire, car elle permettra de révéler des caractéristiques propres à l'usage des verbes d'indication de point de vue qui peuvent ne pas apparaître au moyen des descriptions strictement qualitatives.

3.4.1 Fondements théoriques de l'analyse cooccurentielle

Notre approche puise dans la tradition de l'école contextualiste britannique, mais aussi des plus récents développements de la textométrie française. Les travaux de Blumenthal (2002, 2012) nous inspirent également en ce qui concerne l'approche comparative. Nous rappelons brièvement les principes de ces écoles et études.

Le courant linguistique du contextualisme britannique, dont les principaux représentants sont J. R. Firth et J. Sinclair, encourage la nécessité de prendre en considération les contextes pour saisir entièrement l'essence sémantique d'un mot :

You shall know a word by the company it keeps. (Firth 1957 : 179)

By far the majority of text is made of the occurrence of common words in common patterns. (Sinclair 1991 : 108).

Comme ces citations le montrent bien, le sens d'un lexème ne se limite pas à sa seule forme, mais s'élabore également à partir des éléments qui l'entourent. Sinclair (1996) parle notamment de « extended units of meaning » (cf. également Tognini-Bonelli 2001, Rundell 2018) pour désigner les mots qui apparaissent soit dans le cotexte direct du lexème en question, soit dans son contexte plus large et qui participent activement à la construction de son sémantisme. On désigne un mot apparaissant dans le même contexte que le lexème en question un **cooccurent** si un lien linguistique entre les deux peut être établi. Si entre ces deux formes il existe une unité syntaxique et sémantique forte, nous sommes en présence d'une **collocation** ; si la signification d'une collocation ne peut pas être dérivée de la signification de chaque composante individuelle, on parlera alors d'**expression figée**.

Pour Harris (1954), plus deux mots ont des entourages semblables, plus ils sont proches sémantiquement :

It is possible to measure how similar are the selection approximations of any two words (within various sets of data). If for two elements A and B we obtain almost the same list of particular environments (selection), except that the environment of A always contains some X which never occurs in the environment of B, we say that A and B are (complementary) alternants of each other: e.g. *knife* and *knive* [...]. If A and B have almost identical environments except chiefly for sentences which contain both, we say they are synonyms: *oculist* and *eye-doctor*. If A and B have some environments in common and some not (e.g. *oculist* and *lawyer*) we say that they have different meanings, the amount of meaning difference corresponding roughly to the amount of difference in their environments. (Harris 1954 : 157)

La tendance d'un mot à apparaître dans des schémas textuels spécifiques a été décrite comme « semantic prosody » (Louw 1993, Stewart 2010) ou comme « semantic preference » (Sinclair 1999). La différence entre les deux n'étant pas évidente (cf. Partington 2004), nous pouvons plus généralement nous référer à ce phénomène dans les termes de Hoey (2006) en parlant de « semantic association » :

The definition of semantic association that we have arrived at is that it exists when a word or word sequence is associated in the mind of a language user with a semantic set or class, some members of which are also collocates for that user. (Hoey 2006 : 24)

Les associations sémantiques sont, pour Hoey (2006), innées dans l'esprit des utilisateur·trices de la langue : lorsque nous entendons ou utilisons un mot, notre esprit établit des connexions avec d'autres mots ou concepts qui partagent une signification ou un contexte similaire. Ces connexions sont basées sur l'expérience linguistique, le contexte culturel et l'utilisation de la langue propres à chacun. Le processus grâce auquel un mot se charge cumulativement avec les cotextes et les contextes dans lesquels il apparaît et le fait que notre connaissance même du mot inclut le fait qu'il apparaît avec certains mots dans certains types de textes est défini par Hoey comme « priming » (Hoey 2006 : 7).

Dans le contexte francophone, les méthodes de la statistique textuelle remontent aux années soixante et septante (Guiraud 1960 ; Muller 1977) et ont été largement influencés par les apports du mathématicien J.-P. Benzécri (Benzécri 1966, 1982 ; Benzécri et al. 1981). L'étude de Lebart et Salem (1994) devient un ouvrage de référence dans ce domaine. Selon l'approche que l'on adopte, on parle aujourd'hui d'analyse textuelle des données, de lexicométrie, de lexicologie, ou encore, plus récemment, de textométrie :

D'abord connue sous le nom de lexicométrie, la démarche a été renommée textométrie (ou logométrie) au début des années 2000 pour rendre compte du fait qu'elle allait au-delà de la seule analyse du lexique, en considérant aussi la morphosyntaxe, l'enchaînement des mots, les structures textuelles et intertextuelles, etc. (Mayaffre et al. 2019 : 102)

La textométrie, centrée sur le texte, a pu être mobilisée par diverses sciences humaines (histoire, littérature, sciences politiques...). Développée au sein d'une communauté scientifique s'intéressant à l'analyse des données textuelles, elle se caractérise notamment par certains calculs fondateurs, statistiques (les spécificités, les cooccurrences) ou non (les segments répétés, les concordances), et accorde une place fondamentale au « retour au texte » (bien outillé dans les logiciels) pour interpréter les unités (généralement des mots) sélectionnées par les calculs. (Pincemin 2011 : 260)

La méthode textométrique ne s'arrête pas à la sélection statistique de mots, elle affirme l'importance de les interpréter en contexte et prévoit plusieurs modes complémentaires de consultation de ceux-ci. (Pincemin 2022 : 379)

La textométrie se distingue d'une approche de linguistique de corpus « à la britannique » car elle insiste sur la nécessité d'appréhender le texte analysé dans sa totalité, en allant au-delà de l'analyse simple du lexique en y incluant la morphosyntaxe, la structure des mots et des textes, et les relations intertextuelles. Elle met en avant l'importance des calculs statistiques et de l'interprétation en contexte des mots sélectionnés, en offrant une approche holistique dans le but de caractériser les types de textes étudiés. Cette volonté d'explorer le texte dans sa globalité et de contribuer à sa description est centrale pour l'approche textométrique, 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). De son côté, la linguistique de corpus n'a pas pour but de proposer une catégorisation du texte, mais « s'attache plutôt à l'analyse d'un marqueur linguistique décrit en usage » (Mayaffre et al. 2019 : 102).

Notre démarche relève ainsi à la fois de la linguistique de corpus et de la textométrie, car les deux méthodologies ne sont pas imperméables et peuvent se combiner, comme le rappelle Pincemin (2011) :

Une étude qui utilise une approche et des outils textométriques, et qui vise à observer et décrire des phénomènes linguistiques en corpus, peut donc relever à la fois de la linguistique de corpus et de la textométrie, sans pour autant qu'aucun de ces deux courants ne subsume l'autre. (Pincemin 2011 : 260).

Toutefois, il est important de rappeler que notre objectif premier n'est pas de décrire les types de discours composant les corpus de travail, mais de décrire les formes sélectionnées tout en nous appuyant sur différents types de textes, dans le but de voir si des régularités d'usage peuvent être mises en relief. Ces régularités pourraient donner des indications sur la codification de traits de sens intrinsèques à la sémantique de ces formes.

Les travaux menés dans la cadre du projet FNS intitulé « L'ancrage argumentatif des formes modales. Etude sur corpus avec un éclairage comparatif entre français et italien », qui donnent une représentation synchronique de la façon dont les formes modales se comportent et influencent les patterns argumentatifs dans lesquels elles apparaissent, constituent le fondement théorique et méthodologique de notre étude. Les recherches qui ont été menées pour arriver à profiler une description fine des certaines formes argumentatives (notamment les verbes modaux et les connecteurs) nous ont inspirée et aidée à fournir ce travail sur les verbes d'indication de point de vue.

Enfin, l'approche contrastive (entre formes et entre langues) de cette étude s'appuie sur les travaux de Blumenthal (2002, 2012), pour qui il est possible d'affiner l'analyse de formes supposées synonymes dans une langue à travers une étude méticuleuse de leur contexte. L'approche contrastive permet, de plus, « d'approfondir la caractérisation des valeurs d'un mot face à ses équivalents supposés dans une autre langue, afin de préciser les contrastes interlinguistiques [...] avec une rigueur que ne permettent les méthodes conventionnelles » (Blumenthal 2002 : 115). La démarche méthodologique de Blumenthal (2012) consiste en trois phases ainsi présentées :

- 1) d'abord, nous observons l'usage des mots à étudier dans des corpus électroniques, parfois à l'aide d'outils informatiques ;
- 2) ensuite, de ces constatations empiriques, nous essayerons de conclure à des éléments du contenu du mot – conformément à l'hypothèse wittgensteinienne qu'usage et signification se répondent ;

3) enfin, ces éléments de contenu, étant en principe de nature abstraite et indépendants d'une langue donnée, sont susceptibles de servir de base pour une comparaison interlinguistique. (Blumenthal 2012 : 113)

Notre façon de procéder est inspirée par ces trois phases. L'observation de l'usage des verbes français dans nos corpus à travers l'étude de ceux que nous pouvons appeler leurs **cooccurents spécifiques (ou cooccurents statistiquement significatifs)**³⁸, nous aide à dresser leur « profil combinatoire » (Blumenthal 2002), à savoir l'image que les cooccurents spécifiques donnent du comportement du verbe en question. Cela permettra de mieux comprendre leur sémantisme et de voir ce qui les distingue entre eux, si tel est le cas. Nous appliquerons la même démarche pour l'italien, avant de procéder à une comparaison interlinguistique français-italien. Dans la section suivante, nous détaillons plus précisément les étapes de notre méthodologie et les outils utilisés pour mener à bien ces analyses.

3.4.2 Choix méthodologiques pour l'établissement du profil combinatoire des verbes d'indication de point de vue

Dans cette section, nous expliquons les démarches qui nous permettent de définir ce que nous appelons, à la suite de Blumenthal (2002), le « profil combinatoire » des verbes d'indication de point de vue, à savoir la « structure schématique du voisinage syntaxique et sémantique d'un mot pivot telle qu'elle se manifeste dans un vaste corpus » (Blumenthal 2002 : 115). Nous articulons l'établissement du profil combinatoire de chaque verbe en plusieurs étapes : après une étape préliminaire de définition des empan pour l'extraction des cooccurents spécifiques (désormais CS), notre méthodologie prévoit une analyse en trois niveaux : un niveau macro (sélection des CS), un niveau intermédiaire (classement des CS) et un niveau micro (analyse des CS et de leurs contextes d'apparition). Cette méthodologie s'applique de la même manière au paradigme des verbes français qu'à celui des verbes italiens. Les recherches ont été menées sur les corpus des discours présidentiels (PRES-FR et PRES-IT) et sur le corpus composé d'articles de presse nationale (LM et REP).

³⁸ Par cooccurrent spécifique ou statistiquement significatif, nous entendons le cooccurrent dont l'association avec le verbe d'indication de point de vue est suffisamment forte pour qu'elle ne soit pas due au hasard, selon un critère probabiliste et un seuil prédéfini de signification statistique.

3.4.2.1 Etape préliminaire : définition des empan et extraction des CS

Pour l'établissement du profil combinatoire des verbes d'indication de point de vue, nous avons choisi d'interroger un voisinage sémantique et syntaxique restreint, en ne prenant en compte que les CS qui interviennent dans la proposition Q (celle qui contient le verbe d'indication de point de vue) et ceux qui se trouvent dans la proposition subordonnée p (la complétive introduite par le verbe). Afin de déterminer la taille d'empan optimale à interroger pour que les CS extraits se trouvent bien soit dans Q soit dans p , plusieurs tests sont réalisés, de manière distincte en français et en italien. Ces tests font varier les empan interrogés et consistent à évaluer la pertinence des listes de CS en fonction de la fenêtre de mots choisie. Ainsi, des empan communs aux deux langues ont pu être fixés, bien que cette convergence ne soit pas évidente. En effet, malgré leur proximité typologique et génétique, le français et l'italien peuvent présenter des fonctionnements divergents, ce qui doit amener le·la chercheur·euse à s'interroger sur la prise en compte d'empan potentiellement différents selon les formes objet d'analyse (cf. par exemple les choix opérés dans Rossari et al. 2024). Pour le contexte gauche, un empan restreint de **4 items** a été fixé dans le but de se focaliser uniquement sur les éléments pouvant venir en appui au verbe pour l'introduction de la complétive, en évitant au maximum ce qui apparaît dans la proposition précédente. Le contexte droit a été élargi à **10 items** dans le but de rendre compte de la manière la plus complète possible des éléments apparaissant dans l'entièreté de la complétive, qui peut être plus ou moins étendue.

A partir des pivots des verbes d'indication de point de vue (cf. la colonne « Paramétrage de la recherche dans TXM » dans les Tableau 2 et Tableau 3), **tous les cooccurrents spécifiques avec un indice de cooccurrence de minimum 3 ont été extraits**, de manière séparée à gauche et à droite du verbe. Cela signifie que, pour chaque verbe d'indication de point de vue, nous avons obtenu deux listes de cooccurrents spécifiques : une présentant tous les CS apparaissant dans l'empan gauche de 4 items, et une affichant tous les CS dans un empan droit de 10 items. Ces listes sont reportées dans les Annexes, section 9.3 : chaque cooccurrent spécifique apparaît sous forme de lemme et est accompagné par la catégorie morphosyntaxique correspondante. A partir de ces listes « brutes », non-triées, s'impose à présent une réflexion sur le type de cooccurrents spécifiques qu'il est pertinent de prendre en considération pour l'établissement du profil combinatoire des verbes d'indication de point de vue.

3.4.2.2 Niveau macro : sélection des cooccurrents spécifiques

Le niveau macro constitue le premier stade de notre analyse. A partir de tous les cooccurrents spécifiques extraits de façon automatique, il est nécessaire de procéder à un tri des listes : cette section présente les réflexions théoriques qui ont orienté notre sélection et les étapes qui ont mené à la visualisation de ces premiers résultats sous forme d'échelles de mise en scène énonciative.

Réflexions théoriques pour la sélection des CS

Les réflexions qui motivent le choix des cooccurrents spécifiques les plus pertinents pour l'analyse des verbes d'indication de point de vue s'ancrent dans le cadre théorique de la pragmatique intégrée de Ducrot, dans lequel nous nous inscrivons. Ducrot théorise le principe que tout énoncé véhicule systématiquement une image de sa propre énonciation :

Le sens d'un énoncé, c'est, pour moi, une description, une représentation qu'il apporte de son énonciation, une image de l'événement historique constitué par l'apparition de l'énoncé [...]. (Ducrot et al. 1980 : 34)

L'idée fondamentale, je le répète, est que tout énoncé, fût-il en apparence tout à fait « objectif » (*La terre est ronde*), fait allusion à son énonciation : dès qu'on parle, on parle de sa parole. Cette hypothèse que le dit dénonce le dire, même lorsqu'il ne le fait pas à l'aide de tel ou tel morphème explicite, cette hypothèse peut jouer un grand rôle dans la description de ces enchaînements d'énoncés qui constituent le discours. (Ducrot et al. 1980 : 40)

[L]e structuralisme dont j'ai parlé ne consiste pas seulement à dire que les énonciations se font allusion les unes aux autres, mais à dire que chacune se fait allusion à elle-même, et au « monde idéal » issu de sa propre existence. Le concept central, ici, c'est celui de sui-référence. [...] (Ducrot 1984 : 85)

Le point fondamental qui nous guide est précisément ce pouvoir que le discours a de « se prendre pour objet, et de se commenter au fur et à mesure de son déroulement » (Ducrot 1984 : 85), en particulier le fait que le locuteur exploite la langue pour qualifier son énonciation. Comme Ducrot, nous considérons que le locuteur est présent dans chaque énoncé. Toutefois, nous nous éloignons de sa perspective en affirmant que certains mots explicitent la voix du locuteur, alors que d'autres ne le font pas. Nous opérons en effet une distinction entre l'énoncé *La terre est ronde* et un énoncé tel que *La terre est belle* : dans le premier, la présence du

locuteur n'est pas explicite, alors que dans le second, elle l'est en raison de l'utilisation de l'adjectif *belle*. Ainsi, pour nous, certaines formes rendent la présence de L dans son énoncé aussi explicite que ne le font les verbes d'indication de point de vue, où L est explicité par le *je* (ou le morphème de première personne pour l'italien). **Seuls les cooccurents spécifiques qui ont ce pouvoir de mettre en scène de manière explicite la voix du locuteur dans son énoncé ont été retenus pour notre analyse.**

Pour illustrer concrètement notre méthode de sélection, fournissons quelques exemples. Prenons les listes de tous les CS de *je crois que* et de *credo che* dans le corpus des discours présidentiels (cf. Annexes, section 9.3). Nous avons retenu des adjectifs qui portent un trait évaluatif tels que *bon* et *important*, ainsi que des adverbes énonciatifs comme *francamente* (franchement) : ce sont des exemples de marqueurs avec lesquels le locuteur manifeste explicitement sa présence au sein de l'énoncé. En revanche, des adjectifs comme *français* ou des noms tels que *démocratie*, *fait*, *parte* (partie), *sanitario* (médecin) n'ont pas été sélectionnés. Il est évident que, afin d'opérer cette sélection, nous ne nous sommes pas simplement fiées aux listes de CS, mais nous avons examiné les contextes d'utilisation de chaque élément afin de vérifier la pertinence des CS retenus. Prenons maintenant la liste des CS de *credo che* dans l'empan gauche dans REP. Nous observons l'adverbe *veramente*, qui peut, selon les contextes, être utilisé en tant qu'adverbe énonciatif (*veramente, credo che si sia sbagliata!*). Cependant, nous remarquons que dans la majorité des cooccurrences, il est employé en tant qu'adverbe de degré et n'apparaît pas dans *Q*, mais dans une proposition antérieure :

(62) Mi sento veramente bene e **credo che** tutti si sentano così. (REP)

Je me sens vraiment bien et je crois que c'est le cas pour tout le monde.

(63) Dobbiamo dire bravo al pittore perché la tonalità di colore era veramente azzeccata. Non **credo che** si possano attribuire particolari responsabilità a Moratti e Galliani, il problema del campo di San Siro ha ormai vent'anni, da quando in occasione dei Mondiali del '90 fu costruito il tetto allo stadio. (REP)

Nous devons dire bravo au peintre car la palette de couleurs était vraiment parfaite. Je ne crois pas que l'on puisse attribuer de responsabilités particulières à Moratti et à Galliani, le problème du terrain de San Siro date désormais d'il y a vingt ans, depuis la construction du toit du stade lors de la Coupe du Monde de 1990.

Nous observons également un autre problème dans (63), à savoir qu'en italien, certains CS interviennent avec des *credo che* qui sont niés. Il est donc crucial de ne sélectionner que les éléments pertinents de la liste (les CS présentant plus de 20% de cooccurrences non pertinentes n'ont pas été retenus). Ainsi, bien que les tests d'empans réalisés lors de l'étape préliminaire permettent de réduire la marge d'erreur, des scories subsistent et il est possible que tout ce qui a été extrait ne soit pas pertinent. Cela souligne l'importance d'un retour systématique au texte tout au long de la recherche.

Création d'une macro-forme de CS

Une fois le tri effectué, nous souhaitons calculer, pour chaque verbe, son attraction globale avec l'ensemble de ses CS sélectionnés. Cela nous permet de déterminer la sensibilité de chaque verbe à ce type de formes. L'objectif est d'examiner, à ce premier niveau, s'il existe une différence entre les verbes d'indication de point de vue et leur capacité à apparaître dans des contextes où se trouvent d'autres marqueurs explicitant la présence du locuteur dans son énoncé.

Pour ce faire, nous regroupons l'ensemble des CS sélectionnés pour chaque verbe (qu'ils apparaissent dans l'empan gauche ou dans l'empan droit) sous une seule et unique macro-forme. Nous souhaitons ainsi calculer la force d'attraction entre chaque verbe et sa macro-forme correspondante : il faut donc imaginer la macro-forme de chaque verbe comme un macro-cooccurent et recalculer l'indice d'attraction statistique entre ce macro-cooccurent et le verbe correspondant. Étant donné que TXM ne permet pas cette manipulation de données, nous utilisons un script dans RStudio pour obtenir la force d'attraction sous forme de *log-likelihood*.

Le *log-likelihood* (cf. Evert 2007) est un indice de probabilité basé sur le même principe que l'indice de cooccurrence de Lafon (1981). Si la fréquence observée d'une forme donnée dans un corpus diffère significativement de la fréquence attendue, le *log-likelihood* sera élevé ($\geq 10,83$), indiquant ainsi que la cooccurrence est statistiquement significative. Le seuil de 10,83 correspond à une probabilité de 0,1% que la cooccurrence soit due au hasard³⁹. Pour que le script calcule le *log-likelihood* entre chaque verbe et chaque macro-forme

³⁹ Les seuils de spécificité des deux indices statistiques utilisés dans cette thèse sont comparables : avec un indice de *log-likelihood* de 10,83 et un indice de spécificité de Lafon de 3, les probabilités d'attraction entre pivot et cooccurent ne sont dues au hasard que dans 0,1% des cas.

correspondante, il est nécessaire de créer un tableau de contingence tel que celui reporté dans la Figure 9 :

PIVOT	OCC	f	f1	f2	N	k	n1	n2
je crois que	macro-forme	519	1056	53521	2843577	14	3	1
je pense que	macro-forme	481	878	58510	2843577	14	3	1
je considère que	macro-forme	100	202	31336	2843577	14	3	1
je dis que	macro-forme	38	102	23979	2843577	14	3	1
je suis sûr·e que	macro-forme	6	90	3239	2843577	14	4	1
je trouve que	macro-forme	22	90	8166	2843577	14	3	1
je suis convaincu·e que	macro-forme	0	68	0	2843577	14	4	1
il me semble que	macro-forme	5	55	492	2843577	14	4	1

Figure 9 : exemple de tableau de contingence (ici pour le corpus PRES-FR).

Dans ce tableau apparaissent les données suivantes :

- **Pivot** : les verbes d'indication de point de vue
- **Occ** : le cooccurrent, à savoir la macro-forme regroupant tous les CS sélectionnés pour chaque verbe
- **f** : la fréquence de la cooccurrence, à savoir combien de fois le pivot rencontre la macro-forme
- **f1** : la fréquence du pivot dans le corpus donné
- **f2** : la fréquence totale du cooccurrent dans le corpus donné (comme nous travaillons avec une macro-catégorie, cela correspond à la somme des fréquences de chaque CS)
- **N** : le nombre total de mots du corpus
- **k** : l'empan interrogé (4 à gauche + 10 à droite)
- **n1** : la taille du pivot (*je crois que* = 3 items ; *je suis convaincu·e que* = 4 items)
- **n2** : la taille du cooccurrent (1 item correspondant à 1 macro-forme)

A partir de ces données, le script calcule le *log-likelihood* en exploitant plusieurs fonctions, que nous reportons dans la section 9.4 des Annexes. Les indices de *log-likelihood* ainsi obtenus permettent de classer les verbes d'indication de point de vue sur une échelle de mise en scène de la voix du locuteur, appelée **échelle de mise en scène énonciative**, et qui visualisera de façon claire la sensibilité de chaque verbe à la macro-forme (cf. sections 4.1 et 5.1).

3.4.2.3 Niveau intermédiaire : classement des CS en catégories sémantiques

Au niveau macro, nous traitons tous les CS ensemble afin de déterminer la sensibilité de chaque verbe d'indication de point de vue à la macro-catégorie de marqueurs explicitant la voix du locuteur. À ce niveau intermédiaire, nous procédons à un classement des CS sélectionnés en fonction de traits sémantiques communs.

Deux classes ont été identifiées pour les CS apparaissant dans l'empan gauche :

- **Construction argumentative** → cette classe regroupe les connecteurs que le locuteur exploite afin de construire une argumentation, comme par exemple *mais/ma* et *donc/dunque*. La voix qui met en lien les contenus et qui signale une prise de position est par défaut celle du locuteur, car ces connecteurs sont étroitement liés au *je* du pivot (ou au suffixe de première personne pour l'italien).
- **Prise en charge** → cette classe regroupe les marqueurs qui indiquent une revendication de la prise en charge du contenu par le locuteur. Cette revendication se fait notamment au moyen de l'explicitation du pronom personnel *moi/io* ou par certains adverbes comme *personnellement/personalmente*.

Dans l'empan droit, deux classes de CS ont également été formées :

- **Modalité** → cette classe regroupe les CS exprimant une modalité du possible ou du nécessaire. Parmi ces marqueurs, nous retrouvons principalement (i) des adverbes et verbes modaux qui ont une interprétation épistémique, comme *probablement, forse* (peut-être), *pouvoir/potere*, au moyen desquels le locuteur se prononce sur les chances de réalisation du contenu de l'énoncé ; (ii) des verbes et périphrases modales qui ont une interprétation déontique, comme *devoir, falloir, être nécessaire/essere necessario* : ces formes permettent au locuteur de donner une indication sur la désidérabilité morale du contenu de l'énoncé (cf. le concept de *morality* évoqué par Nuyts 2016). Dans les deux cas évoqués, c'est la voix de L qui apparaît comme étant responsable de la mise en scène dans l'énoncé de la possibilité ou de la nécessité.
- **Evaluation** → cette classe contient les marqueurs qui sont exploités par le locuteur pour mettre en scène sa réaction émotionnelle, son appréciation ou qualification concernant un état de choses. Pour trier les marqueurs, nous nous basons sur la classification des « subjectivèmes » de Kerbrat-Orecchioni (1980), en ne prenant en compte que les adjectifs et les adverbes subjectifs. Sont considérés comme subjectifs tous les adjectifs et adverbes qui portent un trait sémantique « affectif », comme

vulnérable, antipatico (antipathique), ou un trait « évaluatif », lequel peut être axiologique – *bon/buono, beau/bello* – ou non axiologique – *presque trop grand, competitivo* (compétitif). Nous distinguerons les marqueurs subjectifs des marqueurs objectifs comme *français, parlementaire, secondario* (secondaire), *aggiunto* (ajouté) sans distinguer entre trait « affectif » et trait « évaluatif ».

L'avantage de la perspective intermédiaire par rapport au niveau macro est qu'elle permet de comprendre si une classe sémantique particulière est associée aux verbes d'indication de point de vue. Le même script du niveau macro (cf. section 9.4 des Annexes) sera utilisé pour calculer la force d'attraction entre chaque verbe et les quatre classes de cooccurrents spécifiques. Les résultats de ces analyses seront présentés dans les sections 4.2 et 5.2.

3.4.2.4 Niveau micro : analyse des CS

Le niveau micro présente, sous forme de tableaux (cf. sections 4.3 et 5.3), les CS sélectionnés pour chaque verbe d'indication de point de vue – à savoir uniquement ceux qui mettent explicitement en scène la voix de L. L'objectif de cette troisième étape d'analyse est d'examiner en profondeur la relation entre chaque verbe et ses CS, en se concentrant sur les contextes dans lesquels ils apparaissent. Tandis que les niveaux macro et intermédiaire fournissent des indications respectivement sur l'aptitude de chaque verbe à attirer des marqueurs explicitant la voix de L et sur le type sémantique de ces derniers, le niveau micro permet de mettre en lumière des fonctionnements linguistiques précis. En effet, nous verrons qu'il existe des motifs spécifiques qui émergent de la combinaison de certains verbes avec des cooccurrents spécifiques. La définition de motif que nous adoptons est reprise de Diwersy et Legallois (2014) :

Les motifs ne sont pas de simples configurations syntaxiques ; ce sont plutôt des patrons lexico-grammaticaux (Cellier et al. 2012a et 2012b) qui se définissent par :

- leur récurrence ou leur spécificité ;
- une régularité d'ordre lexical (un même paradigme lexical est employé dans le patron) ou / et fonctionnel : un motif possède une fonction sémantique, pragmatique, rhétorique, voire une fonction d'organisation textuelle.

(Diwersy et Legallois 2014 : 4)

Le repérage de ces motifs dans le niveau micro permet d'observer des similarités et des divergences au sein des verbes d'indication de point de vue d'une même langue, mais aussi entre français et italien.

4. Présentation des résultats de l'analyse cooccurrentielle : les verbes d'indication de point de vue français

Sur la base des études précédentes, nous sommes à présent en mesure de dresser un bref bilan des caractéristiques sémantiques de la majorité des verbes constituant notre paradigme. Les verbes d'indication de point de vue les plus fréquemment étudiés sont *je pense que* et *je crois que*, en raison de leur fréquence d'utilisation : leur fonctionnement est systématiquement comparé du moment que les deux ont la propriété de limiter la validité du contenu de la complétive qu'ils accompagnent et qu'ils sont compatibles avec tout type de prédicat (cf. Blanche-Benveniste et Willems 2007, Gosselin 2015). Cette description convient également à *il me semble que* (cf. Nølle 1994). *Je trouve que* et *je considère que* se différencient de ces trois verbes étant donné leur incompatibilité avec les modalités aléthiques (cf. Gosselin 2015), ou, selon les termes de Ducrot (1975), par le fait qu'ils sont uniquement compatibles avec une prédication originelle.

La méthodologie que nous adoptons dans ce travail permet d'apporter un nouvel éclairage sur les verbes d'indication de point de vue. La dimension quantitative, qui se focalise sur les cooccurrents statistiquement significatifs – un aspect qui n'a jusqu'à présent jamais été abordé –, permet de dessiner un profil combinatoire pour chaque verbe. Par le biais de cette approche, nous aspirons non seulement à confirmer (ou pas) les similarités et distinctions mises en évidence dans les études discutées au chapitre 2, mais également à révéler de nouvelles propriétés qui n'avaient pas encore été identifiées. Comme annoncé dans le chapitre 3, l'analyse cooccurrentielle des verbes d'indication de point de vue s'articulera en trois étapes correspondant à trois niveaux d'analyse : macro, intermédiaire, micro.

4.1 Le niveau macro

Ce premier niveau d'analyse permet d'obtenir une vue d'ensemble sur la propension des verbes d'indication de point de vue à être plus ou moins sensibles aux marqueurs mettant explicitement en scène la voix du locuteur. Dans cette première étape de l'analyse, les CS sélectionnés sont regroupés sous une seule et unique macro-forme, ce qui rendra possible le classement des verbes sur une échelle de mise en scène énonciative.

La Figure 10 permet de représenter visuellement, pour le corpus PRES-FR, le comportement de chaque verbe par rapport à l'ensemble des cooccurrents spécifiques qui ont été retenus. Chaque valeur accompagnant le verbe indique la force d'attraction entre le verbe en question et la macro-catégorie de CS qui lui est associée. Plus la valeur du *log-likelihood* est élevée, plus l'attraction du verbe avec la macro-catégorie est forte. Rappelons que ce n'est qu'à partir du seuil de 10,83 que la valeur du *log-likelihood* indique une association statistiquement significative entre le verbe d'indication de point de vue et la macro-catégorie correspondante ; si cette valeur est inférieure, il n'existe pas d'association statistiquement significative entre les deux.

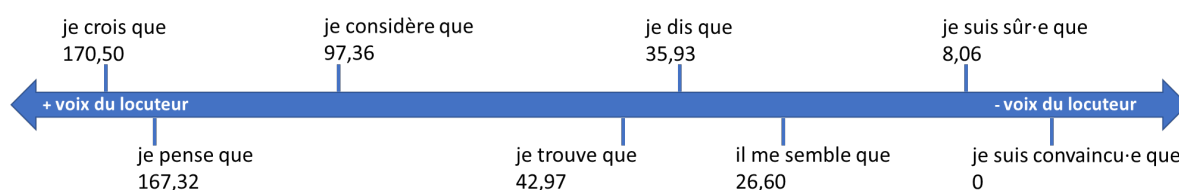


Figure 10 : échelle de mise en scène énonciative des verbes d'indication de point de vue dans PRES-FR.

En observant l'échelle de mise en scène énonciative pour le corpus PRES-FR, il est possible de remarquer que les verbes d'indication de point de vue ne sont pas tous attirés de la même façon par la macro-catégorie des CS. En effet, de grandes différences en termes de valeur du *log-likelihood* sont observables : *je crois que* et *je pense que* présentent des *log-likelihoods* très élevés, respectivement de 170,50 et de 167,32. *Je considère que* apparaît également vers le pôle positif de l'échelle, avec un *log-likelihood* de 97,36. Les autres verbes d'indication de point de vue ont soit des valeurs plus basses (*je trouve que* : 42,97 / *je dis que* : 35,93 / *il me semble que* : 26,60) soit des valeurs qui ne permettent pas d'observer une association statistiquement significative avec la macro-catégorie des CS (*je suis sûr·e que* : 8,06 / *je suis convaincu·e que* : 0). L'échelle pour le corpus LM affiche pratiquement les mêmes tendances que celles présentées pour PRES-FR :

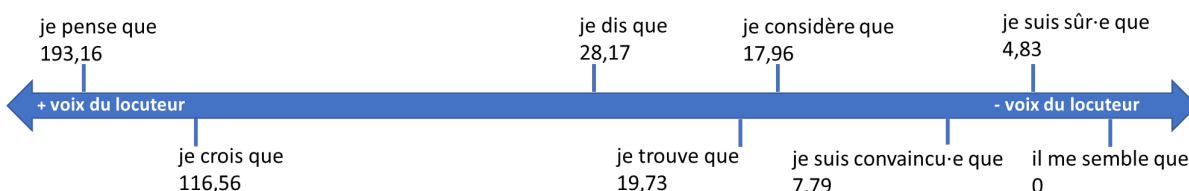


Figure 11 : échelle de mise en scène énonciative des verbes d'indication de point de vue dans LM.

Nous retrouvons, comme dans PRES-FR, *je pense que* et *je crois que* au pôle positif de l'échelle, avec des *log-likelihoods* de 193,16 et de 116,56. De même, *je suis convaincu-e que* et *je suis sûr-e que* se retrouvent au pôle négatif, avec des valeurs (7,79 et 4,83) qui montrent qu'il n'existe pas d'association spécifique entre ces verbes et la macro-catégorie des CS. Dans ce corpus, un troisième verbe ne présente pas d'association spécifique avec la macro-catégorie : il s'agit de *il me semble que*. Les autres verbes d'indication de point de vue (*je dis que*, *je trouve que*, *je considère que*) se trouvent au milieu de l'échelle, avec des valeurs de *log-likelihood* comprises entre 28,17 et 17,96.

4.1.1 Interprétation des échelles de mise en scène énonciative

Les échelles de mise en scène énonciative permettent de montrer des différences entre verbes d'indication de point de vue qui n'auraient pas pu être repérables par une approche strictement qualitative. Les résultats convergents dans les deux corpus montrent qu'une tendance très nette peut être mise en relief entre, d'un côté, *je pense que* et *je crois que* et, de l'autre, *je suis sûr-e que* et *je suis convaincu-e que*. Cette différence en termes d'attraction statistique aux CS mettant en scène la voix du locuteur peut paraître inattendue, du moment que chaque verbe d'indication de point de vue présente *in se* une charge énonciative forte, due à l'ancrage déictique du pronom personnel *je/me*. Du moment que tous les verbes de notre paradigme mettent fortement en jeu la voix du locuteur, on serait amenée à penser qu'il n'y a *a priori* pas de raison que les contextes d'apparition diffèrent. Observons ces deux exemples illustrant les verbes d'indication de point de vue qui présentent une forte attraction à la macro-catégorie des CS dans les deux corpus⁴⁰ :

(64) **Je pense que** la mauvaise situation économique de l'Algérie a dû beaucoup contribuer à accroître le nombre de partisans des mouvements extrémistes. (PRES-FR)

(65) J'ai vraiment une idée très floue du niveau de mes élèves. J'avais préparé une compréhension orale pour une classe de 2de qui a été un

⁴⁰ Dans tous les exemples que nous utiliserons dorénavant, nous fournirons un contexte parfois étendu afin de garantir au·à la lecteur·trice une compréhension optimale. Cependant, rappelons que, conformément aux critères établis dans la section 3.4.2.2 (taille des empan, pertinence des CS), nous avons pris en compte pour notre analyse uniquement les marqueurs soulignés.

fiasco complet. **Je crois que** mon texte était bien trop difficile. Heureusement, ma préparation comportait aussi un travail sur images et j'ai pu rebondir là-dessus [...]. (LM)

Les contextes dans lesquels ces deux verbes apparaissent font transparaître la voix du locuteur à plusieurs reprises : dans (64), le locuteur décrit la situation économique de l'Algérie comme *mauvaise* et formule ensuite des hypothèses au moyen du verbe modal *devoir* et de l'adverbe *beaucoup*. Dans (65), observons également, tant dans le contexte proche que dans le contexte plus large, la présence d'adverbes de degré ou évaluatifs (*vraiment, heureusement*) et la surabondance d'adjectifs et d'adverbes subjectifs (une idée *très floue*, un fiasco *complet*, un texte *bien trop difficile*). Tous ces marqueurs, accompagnés par les verbes respectifs, mettent en avant les impressions et les points de vue du locuteur. La substitution par un autre verbe de notre paradigme n'entraînerait pas, pour aucun des deux exemples, un changement de sens radical des énoncés. Prenons par exemple l'extrait (64) et essayons d'utiliser, à la place de *je pense que*, les autres verbes de notre paradigme :

- (66) **Je crois que** la mauvaise situation économique de l'Algérie a dû beaucoup contribuer à accroître le nombre de partisans des mouvements extrémistes.
- (67) **Je trouve/considère que** la mauvaise situation économique de l'Algérie a dû beaucoup contribuer à accroître le nombre de partisans des mouvements extrémistes.
- (68) **Je suis sûr·e/convaincu·e que** la mauvaise situation économique de l'Algérie a dû beaucoup contribuer à accroître le nombre de partisans des mouvements extrémistes.
- (69) **Je dis que** la mauvaise situation économique de l'Algérie a dû beaucoup contribuer à accroître le nombre de partisans des mouvements extrémistes.
- (70) **Il me semble que** la mauvaise situation économique de l'Algérie a dû beaucoup contribuer à accroître le nombre de partisans des mouvements extrémistes.

Tous ces verbes sont acceptables dans ce contexte car ils permettent tous au locuteur de présenter le même contenu (*la mauvaise situation économique de l'Algérie a dû beaucoup contribuer à accroître le nombre de partisans des mouvements extrémistes*) sur un ton engagé (Carel 2011a). Nous pouvons néanmoins mettre en avant quelques nuances de sens sur la base des études qualitatives présentées dans le chapitre 2. En effet, nous pouvons avancer que le locuteur de (67) a une expérience directe de la situation dont il parle et il exprime un jugement sur *p* qui relève de l'évaluatif. Dans les autres exemples, les différences se dessinent au niveau de la valeur de vérité de *p* : si dans (68) et (69) le locuteur y adhère, dans (70) il l'affaiblit. L'interprétation de l'exemple (66), tout comme celle de (64), reste indéterminée entre une lecture modale, transparente (au moyen de laquelle le locuteur ne fait que poser *p*) et une lecture attributive ou descriptive (lorsque le focus est mis sur le processus psychologique du locuteur). Les échelles que nous venons de présenter apportent des éléments d'analyse complémentaires : elles font ressortir très clairement des tendances d'utilisation de *je pense que* et *je crois que*, qui ont une préférence nette pour les contextes dans lesquels le locuteur montre sa voix en se mettant en scène dans son discours.

Il est toutefois important de noter que, bien que certains verbes d'indication de point de vue montrent une tendance statistique à ne pas avoir d'affinité marquée avec la macro-catégorie des CS, cela n'implique pas pour autant une incompatibilité avec ce type de marqueurs. En effet, des formes comme *je suis sûr-e que* et *je suis convaincu-e que* peuvent également intervenir dans ces contextes, comme le démontrent ces exemples :

- (71) Pourvu qu'au cours de 1965 notre pays maintienne la même stabilité et la même activité, **je suis sûr que** l'année qui commence sera une étape importante de son extraordinaire développement. (PRES-FR)
- (72) Je vous remercie, messieurs, de votre aide, de votre appui, et **je suis convaincu que** cet exemple symbolique sera suivi par d'autres. (PRES-FR)

Dans l'exemple (71) interviennent les adjectifs subjectifs *importante* et *extraordinaire*, dans (72) *symbolique*. En termes de positionnement énonciatif du locuteur, ces deux exemples ne sont pas très différents des exemples (64) et (65), qui mobilisent *je pense/crois que* : dans tous les cas le locuteur se fait entendre via l'utilisation des formes verbales et des adjectifs subjectifs. En synthèse, l'apport fondamental des échelles de mise en scène énonciative est celui de montrer que, malgré le fait que tous les verbes d'indication de point de vue du

paradigme sont compatibles avec des contextes où le locuteur montre sa voix, il existe des tendances statistiques qu'il ne serait pas possible de repérer sans la mise en place de cette méthodologie. Il ne serait en effet pas possible de prédire que *je pense que* et *je crois que* sont beaucoup plus sensibles à des CS que ne le sont *je suis sûr-e que* et *je suis convaincu-e que*.

4.1.2 Synthèse du niveau macro

Ce premier niveau d'analyse a permis d'esquisser une idée générale du comportement des verbes d'indication de point de vue par rapport à la macro-catégorie des marqueurs de mise en scène énonciative qui leur sont spécifiques. Si la majorité des verbes du paradigme présente une tendance statistique à apparaître dans des contextes où le locuteur fait entendre sa voix, les échelles de mise en scène énonciative ont néanmoins mis en relief deux particularités partagées par les corpus PRES-FR et LM. D'un côté, *je pense que* et *je crois que* présentent des indices de *log-likelihood* nettement plus élevés que tous les autres verbes ; de l'autre, *je suis sûr-e que* et *je suis convaincu-e que* ne sont pas spécifiques à la macro-catégorie des CS. Pour approfondir et mieux comprendre ces résultats, nous passons à présent au niveau intermédiaire, qui permettra de mettre en évidence, pour chaque verbe, quelles sont les classes sémantiques qui contribuent majoritairement à la construction de la macro-catégorie dont il a été sujet ici.

4.2 Le niveau intermédiaire

Au niveau macro, les verbes d'indication de point de vue ont été classés sur une échelle de mise en scène énonciative en se basant sur leur indice de *log-likelihood* avec la macro-catégorie des CS. Au niveau intermédiaire, nous pouvons affiner cette analyse en identifiant plus précisément les classes de marqueurs qui contribuent le plus à la construction de cette macro-catégorie. Rappelons que pour chaque verbe, en partant de tous les CS retenus, nous avons pu distinguer quatre classes :

1. La classe **Construction argumentative** comprend les CS grâce auxquels le locuteur peut établir des rapports argumentatifs entre les unités linguistiques de son discours, comme les connecteurs ;
2. La classe **Prise en charge** comprend les CS que le locuteur utilise pour revendiquer le contenu qu'il présente, comme les adverbes d'énonciation et les pronoms personnels ;

3. La classe **Modalité** comprend les CS exprimant une modalité du possible ou du nécessaire, comme les verbes et les adverbess modaux ;
4. La classe **Evaluation** comprend les CS à travers lesquels le locuteur exprime son état appréciatif ou émotionnel sur un contenu donné, comme les adjectifs et les adverbess subjectifs.

Dans les Annexes, nous reportons un tableau récapitulatif (Tableau 9) qui a pour but de donner une vision globale de tous les CS que nous avons rencontrés, et qui peuvent contribuer, selon le verbe, à la construction des classes sémantiques. Rappelons également que les CS des classes 1 et 2 apparaissent dans l'empan gauche des verbes, les CS des classes 3 et 4 dans l'empan droit. La force d'attraction entre le verbe et chacune des catégories est marquée par l'indice du *log-likelihood*, comme pour le niveau macro. Nous présentons ci-dessous les graphiques indiquant les valeurs du *log-likelihood* de chaque classe sémantique. Pour chaque corpus, les verbes d'indication de point de vue apparaissent dans le même ordre que celui présenté dans les échelles de mise en scène énonciative (plus forte présence de la voix du locuteur à gauche, moins forte présence de la voix du locuteur à droite).

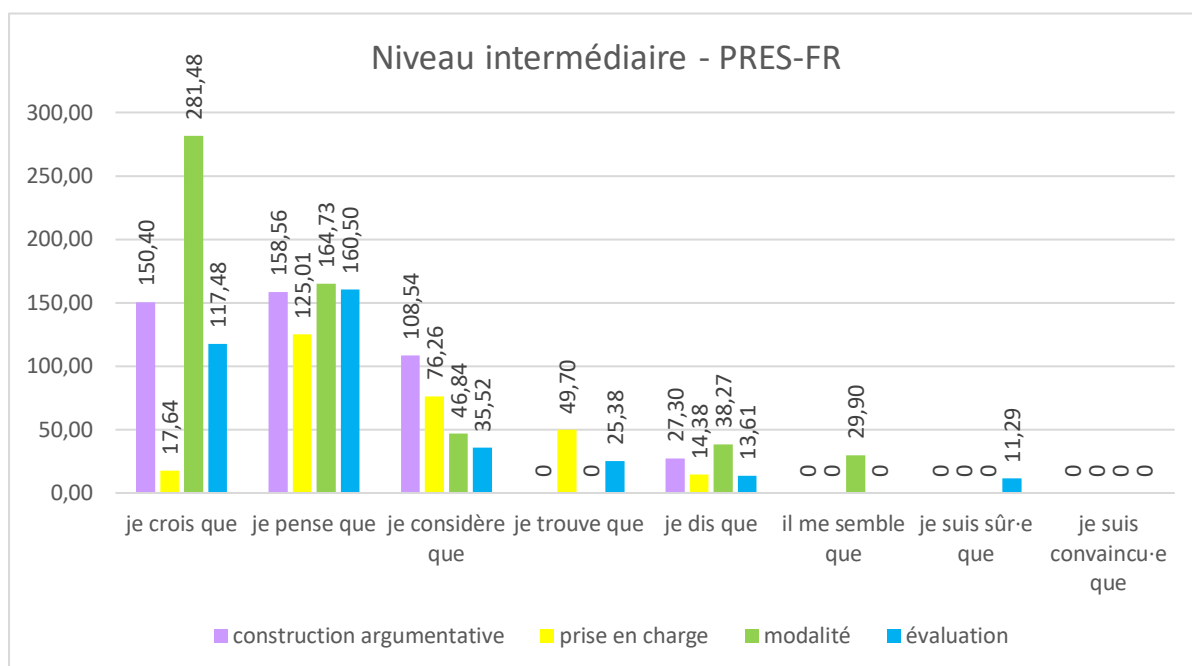


Figure 12 : spécificités entre les verbes d'indication de point de vue et les classes sémantiques des CS dans PRES-FR.

Pour le corpus PRES-FR, nous remarquons d'emblée que les verbes d'indication de point de vue ayant été présentés comme les plus énonciatifs dans la section 4.1 (*je crois que, je pense que, je considère que*) sont sensibles aux quatre classes, avec des valeurs de *log-likelihood* en

règle générale très élevées. *Je dis que* est également associé aux quatre classes, bien qu’avec des valeurs de *log-likelihood* inférieures. Les deux verbes qui ne présentaient pas d’association spécifique avec la macro-catégorie dans le niveau macro, à savoir *je suis sûr-e que* et *je suis convaincu-e que*, montrent respectivement seulement une légère attraction pour la classe *Evaluation* (11,29) et aucune attraction aux quatre classes. Enfin, *je trouve que* est spécifique aux classes *Prise en charge* (49,70) et *Evaluation* (25,38) et *il me semble que* uniquement à la classe *Modalité* (29,90). Passons maintenant au corpus LM :

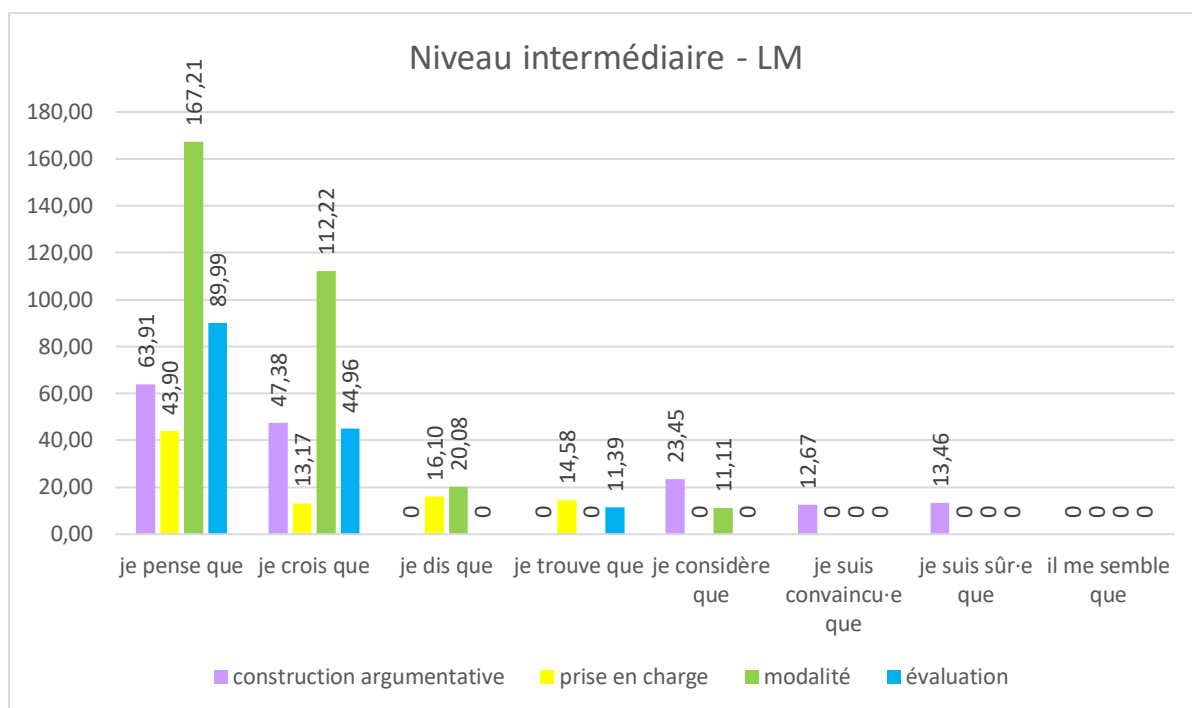


Figure 13 : spécificités entre les verbes d’indication de point de vue et les classes sémantiques des CS dans LM.

Ce corpus présente les mêmes tendances que PRES-FR. En effet, *je pense que* et *je crois que* sont spécifiques aux quatre classes, avec des valeurs de *log-likelihood* un peu inférieures à celles dans PRES-FR, mais quand même très élevées. Les verbes qui n’atteignaient pas le seuil de spécificité avec la macro-catégorie dans le niveau macro ne sont sensibles à aucune classe (*il me semble que*) ou à seulement une classe (*je suis sûr-e que*, *je suis convaincu-e que*). Les verbes occupant une position intermédiaire dans l’échelle présentent chacun une attraction à deux classes sémantiques : *je dis que* aux classes *Prise en charge* et *Modalité*, *je trouve que* aux classes *Prise en charge* et *Evaluation* et *je considère que* aux classes *Construction argumentative* et *Modalité*.

4.2.1 Interprétation des graphiques du niveau intermédiaire

Les graphiques que nous venons de présenter permettent de mettre en relief un procédé qui marque encore une fois une différence entre les verbes se trouvant aux rangs diamétralement opposés sur l'échelle de mise en scène énonciative :

- **Les verbes qui occupent les rangs les plus élevés présentent des associations spécifiques avec toutes les classes sémantiques.**

Je pense que et *je crois que* montrent une tendance très nette à attirer les quatre classes, sans qu'une en particulier ne domine, dans les deux corpus. Le niveau intermédiaire permet non seulement de mettre en relief un comportement extrêmement stable de ces deux verbes dans deux corpus différents, en offrant une vue plus complète sur la composition de la macro-catégorie, mais également de confirmer la pertinence de l'analyse du niveau macro. En effet, il est possible d'exclure la possibilité que les indices d'association très forts présentés dans 4.1 soient dus à des idiosyncrasies qui lieraient l'usage d'un verbe d'indication de point de vue à un marqueur particulier, en formant des unités lexicales bien précises. Le fait que les quatre classes soient équitablement représentées, et que les marqueurs qui les composent soient très diversifiés – comme nous le verrons dans 4.3 –, permet d'exclure la possibilité que les indices élevés soient la cause d'un motif linguistique précis en jeu⁴¹. Cela démontre que leur hypersensibilité aux CS n'est pas le fruit d'une combinaison hasardeuse, mais que les deux verbes ont un fonctionnement très particulier et stable, qui découle d'un processus profond commun à divers types de corpus.

- **Les verbes qui se trouvent aux rangs inférieurs ont tendance à effectuer des sélections précises d'une classe ou deux.**

Les deux verbes qui ne montrent pas d'affinité avec la macro-catégorie des CS dans les deux corpus (*je suis sûr-e que* et *je suis convaincu-e que*) présentent soit une absence totale d'attraction envers les classes sémantiques, soit une sélection beaucoup plus restreinte. Dans le corpus PRES-FR, cette sélection est relativement diversifiée (*je suis sûr-e que* est associé à la classe *Évaluation*, tandis que *je suis convaincu-e que* n'est lié à aucune classe spécifique). En revanche, dans le corpus LM, la situation est plus

⁴¹ Pensons notamment aux cas de collocations qui relèvent des expressions figées.

cohérente, les deux expressions étant spécifiquement associées à la classe *Construction argumentative*. Cette observation révèle une différence dans leur comportement en fonction du contexte discursif, suggérant que ces deux verbes sont davantage ancrés dans des structures argumentatives spécifiques dans le corpus journalistique.

4.2.2 Synthèse du niveau intermédiaire

Le niveau intermédiaire a permis d'identifier les classes sémantiques qui sont majoritairement responsables des indices des macro-catégories présentées dans les échelles de mise en scène énonciative. Les verbes d'indication de point de vue qui se trouvent aux rangs supérieurs de l'échelle (*je pense que* et *je crois que*) sont sensibles à toutes les classes, alors que les verbes qui se trouvent aux rangs inférieurs ont tendance à effectuer des sélections précises d'une ou plusieurs classes. De plus, si certains verbes ont un comportement très stable dans les deux corpus car ils présentent une attraction envers les mêmes classes (*je trouve que*), d'autres semblent avoir moins de points communs (*je dis que*, *je considère que*). Pour mieux comprendre les raisons de ces régularités et différences, nous passons à présent au niveau micro, qui se focalisera sur l'étude des différents marqueurs lexicaux venant construire les quatre classes sémantiques.

4.3 Le niveau micro

Ce troisième niveau d'analyse apporte un degré de granularité accru : en se concentrant plus précisément sur le lexique, il clarifie la composition des classes sémantiques en identifiant les cooccurrents spécifiques qui les constituent. Grâce à ce niveau d'analyse, il est possible non seulement de visualiser les items lexicaux propres à chaque verbe dans leur contextes gauche et droit, mais aussi d'évaluer dans quelle mesure ces marqueurs sont communs ou exclusifs parmi les différents verbes d'indication de point de vue.

Les résultats sont présentés pour chaque corpus dans des tableaux distincts. Les verbes d'indication de point de vue se trouvent au centre, classés selon leur place sur l'échelle de mise en scène énonciative du niveau macro, des plus énonciatifs en haut aux moins énonciatifs en bas. Les deux classes sémantiques des colonnes de gauche (*Construction argumentative* et *Prise en charge*) répertorient les cooccurrents spécifiques avec chaque verbe dans l'empan gauche (4 items). Les classes des colonnes de droite (*Modalité* et *Évaluation*) présentent les

cooccurents spécifiques avec les verbes dans l’empan droit (10 items). Tous les cooccurents spécifiques sont reportés sous forme de lemmes, accompagnés de leur catégorie morphosyntaxique⁴² et de l’indice de spécificité, qui mesure leur force d’attraction avec le verbe correspondant. Il convient de rappeler que c’est l’indice de Lafon (1981) qui est utilisé pour ce niveau d’analyse : avec un seuil minimum de 3, la cooccurrence est statistiquement significative. Nous précisons enfin que les tableaux sont des tableaux récapitulatifs : le fait que les cooccurents spécifiques des différentes classes apparaissent côte à côte ne signifie pas qu’ils apparaissent simultanément dans les cotextes étudiés. Ces tableaux visent simplement à résumer, de la façon la plus lisible possible, les listes de tous les cooccurents spécifiques (CS) retenus pour chaque verbe. La Figure 14 présente les résultats pour le corpus PRES-FR.

CONSTRUCTION ARGUMENTATIVE		PRISE EN CHARGE			MODALITE		EVALUATION	
marqueur	indice de spécificité	marqueur	indice de spécificité		marqueur	indice de spécificité	marqueur	indice de spécificité
KON_parce que	21,46	PRO:PER_moi	4,71	JE CROIS QUE	VER_falloir	41,78	ADJ_bon	7,59
KON_mais	15,67				VER_pouvoir	10,96	ADV_très	7,41
					ADJ_nécessaire	5,80	ADJ_important	6,60
					VER_devoir	4,31	ADV_exagérément	5,51
							ADJ_véridique	4,86
						ADV_surtout	3,52	
						ADV_peu	3,27	
KON_mais	17,82	PRO:PER_moi	18,35	JE PENSE QUE	VER_falloir	19,87	ADJ_bon	12,75
KON_parce que	13,29	ADV_personnellement	15,58		VER_devoir	13,57	ADJ_important	6,74
KON_donc	6,76				VER_pouvoir	6,50	ADJ_sage	6,52
							ADJ_utile	5,39
							ADV_assez	4,41
							ADV_très	3,77
							ADJraisonnable	3,49
						ADJ_urgent	3,46	
						ADVdavantage	3,41	
						ADJ_intelligent	3,17	
KON_parce que	23,21	PRO:PER_moi	15,34	JE CONSIDERE QUE	NOM_devoir	9,59	ADJ_fondamental	3,83
KON_mais	4,85	ADV_personnellement	3,00		VER_devoir	5,33	ADJ_primordial	3,61
KON_car	3,07					ADJ_normal	3,16	
		PRO:PER_moi	9,10	JE TROUVE QUE			ADV_très	6,41
		ADV_personnellement	3,70					
KON_mais	6,82	PRO:PER_moi	3,84	JE DIS QUE	VER_falloir	8,81	ADJ_mauvais	3,59
				IL ME SEMBLE QUE	VER_devoir	4,60		
				JE SUIS SÛR·E QUE			ADV_beaucoup	3,18
				JE SUIS CONVAINCU·E QUE				

Figure 14 : cooccurents spécifiques retenus pour chaque verbe d’indication de point de vue dans PRES-FR.

Les verbes des trois premiers rangs de l’échelle de mise en scène énonciative de PRES-FR (*je crois que*, *je pense que* et *je considère que*) présentent presque systématiquement au moins

⁴² ADJ = adjectif ; ADV = adverbe ; KON = connecteur ; NOM = nom ; PRO:PER = pronom personnel ; VER = verbe.

deux marqueurs différents par classe, avec une abondance importante pour les marqueurs de la classe *Evaluation* pour *je crois que* (7 CS) et *je pense que* (10 CS). Notons d'ailleurs que *je crois que* et *je pense que* partagent, pour chaque classe sémantique, un voire plusieurs marqueurs : *parce que* et *mais* pour la classe *Construction argumentative*, *moi* pour la classe *Prise en charge*, *falloir*, *pouvoir*, *devoir* pour la classe *Modalité* et *bon*, *important*, *très* pour la classe *Evaluation*. *Je considère que* partage également plusieurs marqueurs avec *je crois que* et *je pense que* : *parce que* et *mais* pour la classe *Construction argumentative*, *moi* pour la classe *Prise en charge* et le verbe *devoir* pour la classe *Modalité*.

Les autres verbes ont moins de marqueurs qui leur sont spécifiquement associés : *je trouve que* est caractérisé par l'association spécifique avec deux marqueurs de la classe *Prise en charge* (*moi*, *personnellement*) et un marqueur de la classe *Evaluation* (*très*) ; *je dis que* présente un seul marqueur pour chaque classe (*mais* pour *Construction argumentative*, *moi* pour *Prise en charge*, *falloir* pour *Modalité* et *mauvais* pour *Evaluation*) ; *il me semble que* est uniquement attiré par la classe *Modalité*, représentée par le verbe *devoir*. *Je suis sûr-e que* et *je suis convaincu-e que* – les seuls verbes qui ne présentent pas d'attraction statistiquement significative avec la macro-catégorie des CS – ne sont attirés par aucun marqueur, à l'exception de *beaucoup* (classe *Evaluation*) pour *je suis sûr-e que*, cooccurrence qui a néanmoins un indice de spécificité plutôt faible. Passons maintenant aux résultats pour le corpus journalistique (LM).

CONSTRUCTION ARGUMENTATIVE		PRISE EN CHARGE			MODALITE		EVALUATION	
marqueur	indice de spécificité	marqueur	indice de spécificité		marqueur	indice de spécificité	marqueur	indice de spécificité
KON_mais	11,94	ADV_personnellement	6,31	JE PENSE QUE	VER_falloir	20,83	ADV_assez	6,32
KON_car	3,79	PRO:PER_moi	5,87		VER_pouvoir	9,85	ADJ_bon	6,25
					VER_devoir	6,68	ADJ_préférable	3,83
					ADV_probablement	3,33	ADJ_utile	3,74
							ADJ_noble	3,54
KON_mais	11,28	PRO:PER_moi	3,50	JE CROIS QUE	VER_falloir	13,30	ADV_beaucoup	4,53
					VER_devoir	4,99	ADJ_brave	4,48
					ADJ_possible	4,02	ADJ_tactique	3,63
					VER_pouvoir	3,88	ADJ_vulnérable	3,20
		PRO:PER_moi	4,06	JE DIS QUE	VER_falloir	4,43		
		PRO:PER_moi	3,73	JE TROUVE QUE			ADJ_mauvais	3,03
KON_mais	3,54			JE CONSIDERE QUE	VER_devoir	3,04		
KON_parce que	3,39							
KON_parce que	3,31			JE SUIS CONVAINCU-E QUE				
KON_mais	3,20			JE SUIS SÛR-E QUE				
				IL ME SEMBLE QUE				

Figure 15 : cooccurents spécifiques retenus pour chaque verbe d'indication de point de vue dans LM.

Dans LM, comme dans PRES-FR, *je pense que* et *je crois que* ont une variété de CS beaucoup plus grande par rapport aux autres verbes. Ces deux verbes partagent entre eux le marqueur *mais* pour la classe *Construction argumentative*, le pronom *moi* pour la classe *Prise en charge*, les verbes *falloir*, *devoir*, *pouvoir* pour la classe *Modalité*. Bien que dans ce corpus ils ne partagent pas les mêmes marqueurs pour la classe *Evaluation*, nous retrouvons néanmoins pour les deux tant des adverbes de degré (*assez* pour *je pense que* et *beaucoup* pour *je crois que*), que des adjectifs relevant de l'axiologique (*bon*, *préférable*, *utile*, *noble* pour *je pense que* et *brave* pour *je crois que*). Les autres verbes présentent beaucoup moins de variété de marqueurs, mais nous pouvons tout de même relever une tendance de plusieurs (*je considère que*, *je suis convaincu-e que*, *je suis sûr-e que*) à s'associer à des connecteurs argumentatifs, notamment *mais* et *parce que*.

4.3.1 Points communs entre les deux corpus français

Malgré la diversité des textes composant les deux corpus, des régularités très marquées émergent dans le comportement des verbes d'indication de point de vue. Ces régularités se manifestent dans un premier temps lorsque l'on considère leur charge énonciative au niveau macro, puis leur sensibilité aux différentes classes sémantiques au niveau intermédiaire, et enfin en analysant plus précisément le type de CS avec lesquels ils sont associés. Résumons les points communs que le niveau micro met en évidence entre PRES-FR et LM :

- ***Je crois que*** et ***je pense que*** partagent, dans les deux corpus, les marqueurs *mais* pour la classe *Construction argumentative*, *moi* pour la classe *Prise en charge* et *falloir*, *pouvoir*, *devoir* pour la classe *Modalité*. *Je pense que* partage également les marqueurs *bon*, *utile*, *assez* (*Evaluation*) et l'adverbe *personnellement* (*Prise en charge*).
- ***Je considère que*** présente dans les deux corpus les connecteurs *mais* et *parce que* (*Construction argumentative*) et le verbe *devoir* (*Modalité*).
- ***Je dis que*** partage le pronom *moi* (*Prise en charge*) et le verbe *falloir* (*Modalité*).
- ***Je trouve que*** partage le pronom *moi* (*Prise en charge*).

Le Tableau 4 résume les marqueurs partagés par chaque verbe dans les deux corpus ; le tiret indique l'absence de CS communs et le symbole de l'ensemble vide signifie que l'absence de CS est partagée. Ainsi, lorsqu'une case est verte, le verbe partage les marqueurs indiqués (ou l'absence de marqueurs) dans les deux corpus ; si la case est rouge, le comportement du verbe diffère entre PRES-FR et LM.

<i>Construction argumentative</i>	<i>Prise en charge</i>		<i>Modalité</i>	<i>Evaluation</i>
mais	moi	Je crois que	falloir pouvoir devoir (v.)	-
mais	moi personnellement	Je pense que	falloir pouvoir devoir (v.)	bon utile assez
mais parce que	-	Je considère que	devoir (v.)	-
∅	moi	Je trouve que	∅	-
-	moi	Je dis que	falloir	-
∅	∅	Il me semble que	-	∅
-	∅	Je suis sûr·e que	∅	-
-	∅	Je suis convaincu·e que	∅	∅

Tableau 4 : tableau récapitulatif des CS communs aux verbes d'indication de point de vue dans les corpus français.

Ces résultats montrent l'extrême régularité du comportement des verbes d'indication de point de vue : même dans des contextes discursifs différents, ces verbes conservent des caractéristiques stables en termes de cooccurrence avec les CS.

4.3.2 Interprétation des tableaux du niveau micro

Le Tableau 4 met en évidence le fait que les combinatoires des verbes du paradigme français présentent de nombreux points communs dans les deux corpus étudiés. D'une part, certains verbes se révèlent sensibles à plusieurs, voire à l'ensemble des classes sémantiques. Dans ces cas, ils partagent également de nombreux marqueurs identiques : c'est par exemple le cas de *je pense que*, *je crois que*, *je considère que*, *je dis que*. D'autre part, certains verbes tels que *il me semble que* et *je suis convaincu·e que* se distinguent par leur insensibilité presque totale

aux classes sémantiques, n'étant associés à aucun marqueur pour au moins trois classes, et ce, dans les deux corpus. Dans la mesure où tous ces formes expriment explicitement la voix du locuteur dans son discours, rien ne semble a priori justifier une telle disparité de comportement. La section suivante proposera, pour chaque forme, des éléments d'analyse et de réflexion permettant d'éclairer ces tendances.

4.3.2.1 Je crois que et je pense que

La perspective fine du niveau micro met en lumière la grande variété de cooccurrents spécifiques intervenant dans les environnements proches de *je pense que* et *je crois que*. Que ce soient des CS apparaissant dans le contexte gauche pour introduire le verbe (connecteurs, pronoms personnels, adverbes) ou des CS mobilisés dans la complétive (verbes modaux, adjectifs et adverbes subjectifs), tous indiquent que le réseau d'indications de la voix du locuteur est très développé lorsque *je pense que* et *je crois que* sont mobilisés en discours. Le fait que ces deux verbes d'indication de point de vue partagent un grand nombre de marqueurs précis, et ce dans deux corpus composés des textes à visées différentes, est l'indice que le profil combinatoire de *je pense que* et *je crois que* peut pointer des propriétés discursives stables, non sujettes à l'influence du type de texte et donc potentiellement intrinsèques à leur sémantisme même. Les prochaines pages présentent les usages de ces deux expressions en interaction avec les CS des quatre classes.

Les associations avec les CS de la classe *Construction argumentative*

Pour ce qui concerne la classe *Construction argumentative*, *je crois que* et *je pense que* partagent une association statistique forte, dans les deux corpus, avec le connecteur *mais*. Observons ces exemples tirés des deux corpus :

- (73) Je crois que cette forme-là a disparu tout à fait, mais **je crois que** la famille s'est reconstituée autrement, et qu'elle a une force, à l'heure actuelle, très grande chez les jeunes. (PRES-FR)

- (74) C'était une formidable expérience, mais **je crois qu'**on peut aussi arriver à un résultat satisfaisant avec les instruments d'aujourd'hui. (LM)

- (75) Ensuite naturellement, il faudra procéder à un certain nombre d'études, **mais je pense qu'**un premier projet pourrait être présenté au parlement au printemps prochain. (PRES-FR)
- (76) Je ne demande pas un régime unique, **mais je pense qu'**il est nécessaire de parler de convergence des efforts dans le temps. (LM)

Comme nous l'avons montré dans Rossari et al. (2022), *mais* permet non seulement de hiérarchiser les contenus (propriété « focalisatrice »), mais aussi de rendre manifeste la voix qui prend la responsabilité d'un contenu (propriété « énonciative »). Ainsi, dans les exemples (73) à (76), l'utilisation de *mais* met en scène un contraste énonciatif entre un premier contenu (celui qui précède *mais*) et un deuxième contenu (celui qui suit *mais*). Puisque le contenu à retenir pour l'enchaînement argumentatif est celui qui suit *mais* (cf. Ducrot 1978, Roulet et al. 1987, Anscombe 2002), la perspective énonciative du locuteur qui se met en scène au moyen de *je crois/pense que* se voit renforcée. En effet, les contenus qui précèdent *mais* sont concédés et par conséquent relégués à l'arrière-plan, mis à distance, alors que le point de vue personnel du locuteur est revendiqué et par conséquent porté à l'avant-plan.

Les associations avec les CS de la classe *Prise en charge*

Je crois que et *je pense que* partagent également, dans les deux corpus, une attraction statistique avec le marqueur *moi* de la classe *Prise en charge*. Ce pronom vient en appui aux verbes d'indication de point de vue en introduisant la prise de position du locuteur et en la renforçant, comme le montrent les exemples suivants :

- (77) La France grandira au travers de l'Europe, sa voix sera plus entendue. Et pour chaque profession, le fait d'être associée intimement à des professionnels des autres pays devrait normalement permettre à la France de franchir une étape dans l'évolution moderne. **Moi, je crois que** l'Europe est très importante. (PRES-FR)
- (78) Car pour vous, il est préférable de ne rien dire. Mieux vaut garder le silence. Voir et se taire [...]. Exactement ce que veulent les mafias, en somme, que leurs affaires restent leurs affaires, et d'ailleurs Cosa Nostra est d'abord une expression avant d'être le nom d'une organisation. **Moi, je crois que** seule la vérité sert à donner à un pays sa dignité. (LM)

- (79) 60000 caméras seront installées d'ici 2012. Je laisserai ceux qui le veulent crier à l'atteinte aux libertés individuelles. **Moi je pense que** la liberté individuelle est gravement atteinte lorsque les voyous font régner la terreur devant des immeubles d'habitation. (PRES-FR)
- (80) [...] je sais bien que parfois on me moque parce que je vais tout le temps dans les usines mais **moi, je pense qu'**une grande Nation doit savoir fabriquer des autos, des bateaux, des avions, des trains, qu'il y a une tradition industrielle française, que cette tradition industrielle, si on la laisse partir ailleurs, on ne la retrouvera jamais et que, quand les usines et l'industrie seront parties, ensuite on fera partir les services et il n'y aura plus rien. (PRES-FR)

Lorsque le pronom *moi* est utilisé à côté des verbes exprimant une opinion, il a une fonction de *booster* (Detges et Waltereit 2014 : 39), c'est-à-dire qu'il présente l'opinion du locuteur comme particulièrement originale et digne d'intérêt. Ainsi, grâce à ce pronom, le locuteur met le focus sur son propre point de vue en le faisant contraster avec un autre point de vue intervenant dans le contexte précédent⁴³. Ainsi, dans (77), l'opposition se fait entre un point de vue exprimé à travers un ton factuel, sur le mode du « trouvé » (Carel 2011a : 10), à savoir *la France grandira au travers de l'Europe*, et un point de vue exprimé via un ton engagé, celui du locuteur, sur le mode du « conçu » (Carel 2011a : 10). Ce ton engagé est d'abord mis en scène par des marqueurs épistémiques (*devrait, normalement*), à travers lesquels le locuteur présente l'évolution de la France comme une possibilité, et ensuite par la combinaison de *moi* et du verbe *croire*, au moyen desquels le locuteur prend explicitement en charge les propos présentés jusque-là (à savoir le fait que l'Europe est très importante pour la France et que la France doit donc collaborer avec elle). Dans les exemples (78) à (80), le locuteur oppose son propre point de vue (exprimé par *moi je pense/crois que*) à des points de vue externes qu'il ne partage pas. Cela lui permet de critiquer l'« omertà » de certaines personnes (78), de rassurer les citoyens par rapport à l'installation de caméras de surveillance (79) et de se justifier par rapport à certaines critiques (80).

⁴³ Ce comportement est similaire à celui des expressions *pour moi* et *selon moi*, qui « situent le discours de L dans la continuité : tout se passe comme si au moyen de *pour moi* et de *selon moi* L enchaînait sur un discours ou une réflexion en cours ou sur une question existant antérieurement. » (Coltier et Dendale 2004 : 56).

Une distinction peut être mise en relief entre *je crois que* et *je pense que* si l'on prend en considération un autre marqueur de la classe *Prise en charge*, à savoir l'adverbe *personnellement*. En effet, *je pense que* se distingue par une association statistiquement significative, dans les deux corpus, avec cet adverbe, alors que ce n'est pas le cas pour *je crois que*. Lorsqu'il est combiné avec *je pense que*, *personnellement* accentue l'implication du locuteur en soulignant explicitement sa prise de position, comme en témoignent les exemples suivants :

- (81) Personnellement, **je pense que** les formes extrêmes de l'art contemporain rejoignent les formes extrêmes de l'art traditionnel. (LM)
- (82) Personnellement, **je pense que** les bonnes affaires, lorsqu'elles sont proposées au public à un prix modéré, se vendent bien. (PRES-FR)

Dans des exemples (81) et (82), l'adverbe *personnellement* a une fonction que Guimier (1996) définit « paradigmatique contrastive », car il « attire l'intention de l'interlocuteur sur le fait que ce qui va être dit est pris en charge par le locuteur, lequel reconnaît implicitement que tel autre pourra ne pas partager son point de vue » (Guimier 1996 : 156). Le fait que cet adverbe soit un cooccurrent spécifique de *je pense que* mais pas de *je crois que* peut être mis en relation avec l'analyse de Ducrot (1975). Selon lui, seul le sémantisme de *je pense que* implique un certain degré de réflexion de la part du locuteur (critère R). Le *personnellement* renforce cette idée en indiquant que le point de vue exprimé émane directement du locuteur.

Les marqueurs de *Prise en charge* peuvent également se combiner avec les marqueurs de *Construction argumentative*, comme le montrent les exemples suivants :

- (83) Moi, personnellement, **je pense que** dans certains pays – et cela arrive souvent en France –, il y a, à certains moments, quand on vit profondément la vie de notre histoire, une sorte d'élan national. (PRES-FR)
- (84) C'est au gouvernement d'examiner cela : il m'en parlera nécessairement, puisque j'ai une responsabilité particulière dans ce domaine. Mais moi, personnellement, **je pense que** tout ce qui peut être fait – il y a des données économiques évidentes – pour maintenir les bonnes relations actuelles sera une chose utile. (PRES-FR)

Dans ces deux exemples, le recours à la suite (*mais*) + *moi* + *personnellement* + *je pense que* met en relief une prise de position qui est subjective et assumée, et qui renvoie de façon redondante au locuteur. Ces marqueurs favorisent une lecture attributive de *je pense que*, puisque le locuteur devient objet de son propre discours. Si nous reprenons la distinction de Ducrot entre locuteur « L » et locuteur « λ » (cf. section 1.1), c'est λ qui intervient dans ces emplois. En effet, son but n'est pas de présenter le contenu *p*, de le modaliser, mais de qualifier son énonciation en s'attribuant la propriété de « penser que *p* ». Nous observons d'ailleurs que lorsque les marqueurs de *Prise en charge* sont mobilisés, c'est souvent parce que le locuteur a besoin de faire contraster son point de vue avec un autre. Pour s'opposer à ce dernier, il doit se réapproprier son propre discours et se montrer le plus responsable possible de celui-ci. Cette mise en scène de λ, parfois très ostentatoire – cf. l'exemple (84) – ne peut donc qu'appeler une lecture attributive de *je pense que*, car ce qui est sélectionné dans le sémantisme du verbe est la description d'un état psychologique et non sa valeur modale.

Les associations avec les CS de la classe *Evaluation*

En ce qui concerne la classe *Evaluation*, *je pense que* partage dans les deux corpus une attraction statistiquement significative avec deux adjectifs (*bon*, *utile*) et un adverbe (*assez*). Pour *je crois que*, nous n'observons pas les mêmes marqueurs d'un corpus à l'autre. Toutefois, *je crois que* se rapproche beaucoup de *je pense que* dans le corpus des allocutions présidentielles, car les deux premiers CS de la classe *Evaluation* qui apparaissent sont les adjectifs *bon* et *important* ; l'adverbe *très* est également partagé par les deux verbes dans ce corpus. Ces résultats montrent ainsi que l'emploi *je pense que* et *je crois que* est fortement lié à des prédications originelles (cf. Ducrot 1975), où le but du locuteur est de coller une nouvelle étiquette à un objet donné :

(85) Mais, cela dit, je vous remercie de l'occasion que vous m'avez fournie, car une explication et un débat devant un grand nombre de français, **je crois que** c'est une bonne chose, parce que la démocratie c'est aussi le débat. (PRES-FR)

(86) Lorsque le Mozambique ou l'Angola se tournent vers nous pour dire : « On a besoin de vous », lorsque certains pays anglophones viennent dire : « Moi, je voudrais me rattacher à la zone franc », **je pense que** tout cela est très bon pour le français. (PRES-FR)

(87) **Je crois qu'**il est beaucoup plus important que vous ne le pensez, pour une société, d'avoir une direction, et donc d'avoir des intentions. (PRES-FR)

(88) Vous êtes revenus sur quelques propositions très concrètes, mais **je pense que** c'est un élément très important. (PRES-FR)

Les étiquettes mobilisées dans ces exemples (*bonne, bon, important*) correspondent à des adjectifs axiologiques qui « reflètent la compétence idéologique » du locuteur, en manifestant « une prise de position en faveur, ou à l'encontre, de l'objet dénoté » (Kerbrat-Orecchioni 1980 : 91). Quand un jugement axiologique est employé, son rôle en discours est double : il permet une « persuasion orientée vers l'action tout en fournissant une bonne image du locuteur puisqu'il s'inscrit dans un discours du bien/de l'utile/du désirable » (Gosselin 2010 : 351). Souvent, la persuasion visée par le locuteur est appuyée par le recours aux adverbes de degré (par exemple, nous observons dans les exemples (86) à (88) les marqueurs *très, beaucoup, plus*). Ce procédé est d'ailleurs significatif d'un point de vue statistique : il suffit d'observer les marqueurs de la catégorie *Evaluation* pour ces deux verbes, dans les Figure 14 et Figure 15, pour se rendre compte du nombre d'adverbes de degré apparaissant comme cooccurrents droits de *je pense que* et/ou *je crois que* (*très, exagérément, peu, assez, davantage, beaucoup*). En somme, l'étiquetage que le locuteur effectue dans une prédication originelle introduite par *je crois/pense que* correspond, de manière fréquente dans nos corpus, à une séquence de type « adverbe de degré + adjectif axiologique ».

Les associations avec les CS de la classe Modalité

L'analyse de la classe *Modalité* permet de se rendre compte du fait que *je pense que* et *je crois que* sont mobilisés de façon statistique pour introduire des verbes modaux, et présentent des indices particulièrement importants pour ce qui concerne le lemme *falloir*. En effet, ce dernier présente, pour les deux verbes et dans les deux corpus, les indices les plus hauts non seulement de la classe *Modalité*, mais parmi tous les marqueurs des quatre classes (avec *je crois que* : 41,78 dans PRES-FR et 13,30 dans LM ; avec *je pense que* : 19,87 dans PRES-FR et 20,83 dans LM). Nous reportons ci-dessous quelques emplois de *je pense que* et *je crois que* avec ce verbe :

(89) Moi, personnellement **je pense qu'**il faut maintenir le blocus, la coalition internationale pour parvenir à convaincre la Serbie, le cas

échéant la Croatie, le cas échéant les autres, qu'il convient d'en arriver à la discussion pacifique. (PRES-FR)

- (90) J'ai enseigné le latin et le grec, au collège et au lycée. Je me suis battue comme tant d'autres professeurs pour « sauver l'option », mais, depuis près de vingt ans, **je pense qu'il faudrait** sortir du piège de l'enseignement étroit de la langue. (LM)
- (91) S'il arrivait, pour des raisons qui lui appartiennent, qu'il ne soit pas à nouveau candidat, **je pense qu'il faudra** des primaires au sein de l'UMP. (LM)
- (92) Je sais que la question du financement se pose mais **je crois qu'il faut** y aller. (LM)
- (93) **Je crois qu'il faudrait** approfondir notre réflexion sur deux aspects [...]. (PRES-FR)
- (94) C'est le premier ensemble de réformes qui a été engagé et **je crois qu'il faudra** définir, à partir de la rentrée, un second ensemble de réformes. (PRES-FR)

Dans tous ces exemples, *je pense que* et *je crois que* viennent introduire le verbe *falloir* (conjugué soit au présent ou au futur de l'indicatif, soit au conditionnel présent), qui exprime systématiquement des demandes de faire : maintenir le blocus international (89), changer les modalités d'enseignement (90), organiser des élections (91), agir au niveau financier (92), approfondir la réflexion (93), mettre en place des réformes (94). Les deux verbes viennent ainsi personnaliser ces demandes en permettant au locuteur de les prendre personnellement en charge.

Le même type de motif linguistique se vérifie de façon statistiquement significative au moyen du verbe *devoir* (qui apparaît comme cooccurrent spécifique des deux verbes dans les deux corpus) et de l'adjectif *nécessaire* (qui apparaît comme cooccurrent spécifique de *je crois que* dans PRES-FR) :

- (95) **Je pense que** cela devrait être fait aussi dans d'autres cas. (PRES-FR)

- (96) **Je pense que** l'Europe doit se réveiller et comprendre comment le monde est en train de changer. (LM)
- (97) **Je crois que** nous devons absolument continuer de transformer le pays car les résultats sont là. (PRES-FR)
- (98) Je ne parle pas de cela. Je me mets simplement aujourd'hui sur le plan des idées générales, de la considération d'ensemble du sujet. Mais **je crois qu'il est nécessaire** de la préciser. (PRES-FR)

Nous observons que, dans tous les exemples (89) à (98), *je pense que* et *je crois que* sont utilisés pour préparer et anticiper des jugements normatifs exprimés au moyen des verbes modaux *falloir*, *devoir* et de la périphrase modale *il est nécessaire que*. Ces jugements sont tous présentés sur le « mode du conçu » (Carel 2011a), mais l'interprétation de *je pense/crois que* varie selon les marqueurs qui les accompagnent. Dans les exemples (93) et (95), le conditionnel des verbes modaux semble orienter une lecture modale de *je crois/pense que* : le locuteur présente un contenu, qu'il pose en modulant son engagement vis-à-vis de celui-ci. Cependant, le conditionnel n'implique pas nécessairement une lecture modale : dans l'exemple (90), l'adverbe temporel tend à induire une lecture attributive en attribuant au locuteur la propriété de « penser que p » *depuis près de vingt ans*. Dans d'autres exemples, la lecture modale semble plus facilement exclue. Par exemple, dans (89), nous retrouvons la même mise en scène de λ que dans les exemples (83) et (84), au moyen de la dislocation à gauche du pronom *moi* combinée à l'emploi de l'adverbe *personnellement*. Dans (97), la nécessité exprimée à travers le verbe *devoir*, renforcée par l'adverbe *nécessairement*, semble également favoriser une lecture attributive.

Les formes *je pense que* et *je crois que* sont également en situation de cooccurrence statistiquement significative, dans les deux corpus, avec le verbe *pouvoir*, comme l'illustrent les exemples suivants :

- (99) Cela nous donne donc 3 mois mais **je pense qu'on ne peut** pas laisser pendant 3 mois la situation telle qu'elle est. (PRES-FR)
- (100) **Je crois que** l'on ne peut pas ignorer cette montée en puissance d'un fort sentiment d'inégalité : les salariés du privé ont aujourd'hui le sentiment d'être lésés, car ils estiment qu'on leur demande toujours plus d'efforts pour toujours moins de retraite. (LM)

(101) Il y a des problèmes difficiles à résoudre mais **je pense que** nous pouvons faire des progrès. Je suis un optimiste réaliste. Il faut travailler avec le monde tel qu'il est et essayer de le changer. (LM)

(102) **Je crois qu'on peut** aller plus loin et je pense qu'il faut aller plus loin, qu'il faut en même temps tenir compte des difficultés de l'époque. (PRES-FR)

Nous observons que, lorsque *je pense que* et *je crois que* introduisent le verbe *pouvoir*, ce dernier a presque systématiquement une interprétation radicale⁴⁴. Par exemple, dans les deux premiers exemples, l'interprétation de *pouvoir* est du type « interdiction de faire quelque chose » : *on ne devrait pas laisser la situation telle quelle* (99), *on ne devrait pas ignorer le sentiment d'inégalité* (100). Les deux verbes introduisent donc, comme dans les exemples (89) à (98), des jugements normatifs, exprimés cette fois par *pouvoir* à la forme négative. Dans les exemples (101) et (102), l'interprétation du verbe modal est du type « capacité » ou « possibilité matérielle » : *nous sommes en mesure de faire des progrès* (101), *nous sommes en mesure d'aller plus loin* (102). Remarquons comment dans ces deux derniers cas, le fait d'exprimer une capacité conduit, dans la suite du discours, à l'expression des demandes de faire : *nous avons la capacité de faire des progrès, donc il faut travailler pour changer le monde* (101) ; *nous sommes en mesure d'aller plus loin, donc il faut le faire en tenant compte des difficultés* (102).

Le corpus des allocutions présidentielles se caractérise, par rapport au corpus de presse nationale, par un motif linguistique très précis, construit sur le verbe *croire* et non sur le verbe *penser*. En effet, la séquence *je crois que...pouvoir* se révèle être en association statistiquement significative avec le verbe *dire*⁴⁵, ce qui se matérialise par des occurrences comme :

⁴⁴ Etant donné la polysémie de *pouvoir* et *devoir*, un tri systématique de leurs valeurs en cooccurrence avec les verbes d'indication de point de vue a été effectué dans les deux corpus, pour départager les interprétations radicales des interprétations épistémiques (cf. Barbet 2012). Les emplois où l'interprétation épistémique de *devoir* et de *pouvoir* est mobilisée ne dépassent pas le 10% des cooccurrences pour chaque verbe dans chaque corpus. En effet, dans au moins 90% des cas, nous avons des interprétations radicales qui expriment soit une capacité, soit une permission, soit une possibilité matérielle pour *pouvoir* (cf. Le Querler 1996 : 115), soit une obligation théorique ou pratique pour *devoir* (cf. Kronning 1996 : 110).

⁴⁵ Le verbe *dire* apparaît en situation de cooccurrence statistiquement significative (indice Lafon de 24,67) avec la séquence *je crois que...pouvoir* dans un empan droit de cinq items.

- (103) **Je crois qu'on peut dire**, si vous voulez, que d'abord l'Europe a supporté la tempête et ce n'est pas si mal, et ensuite, qu'elle continue à progresser à un rythme que je juge trop lent. (PRES-FR)
- (104) **Je crois que l'on peut dire** aujourd'hui que la situation légale est satisfaisante. (PRES-FR)
- (105) **Je crois qu'on peut le dire** sans vanité, nous avons le meilleur système d'école maternelle au monde [...]. (PRES-FR)

Dans ces trois exemples, l'expression *je crois qu'on peut (le) dire* peut être interprétée comme un type particulier de commentaire métadiscursif (cf. Roulet 1987 : 112) qui spécifie la fonction illocutoire du discours. En effet, elle permet au locuteur de « dire p », mais en même temps de ne pas en assumer complètement la responsabilité, car le fait de « dire p » est non seulement évoqué comme « possible », mais également introduit par *je crois que*, qui subjectivise le tout, en minimisant la prise de responsabilité du locuteur. De plus, la mise en scène d'un ON-locuteur (cf. Anscombe 2010) fait en sorte que la voix du locuteur est noyée parmi celles d'une « communauté plurielle et anonyme » (Anscombe 2010 : 47), lui permettant de se décharger partiellement de la responsabilité de ce qu'il affirme⁴⁶. Remarquons que dans l'exemple (103) s'ajoute également le marqueur discursif *si vous voulez*, qui permet « d'évoquer un décalage entre le dit et le dire et en même temps d'établir une 'consensualité' avec le co-locuteur » (Beeching 2007 : 83). Le fait que cette association statistiquement significative entre la suite *je crois que...pouvoir* et le verbe *dire* apparaisse dans les allocutions présidentielles mais non dans la presse peut s'expliquer par la nature argumentative des textes composant le premier corpus, qui se focalise davantage sur des stratégies rhétoriques de mise en scène de l'ethos discursif de l'orateur.

4.3.2.2 *Je trouve que*

Je trouve que partage avec *je pense que* et *je crois que* une attraction statistiquement significative avec certains marqueurs de la classe *Prise en charge* (*moi* dans les deux corpus et *personnellement* dans PRES-FR), ainsi qu'avec des marqueurs de la classe *Evaluation*. Cette dernière se révèle toutefois moins riche et variée : ce résultat peut sembler plutôt

⁴⁶ Un marqueur très semblable est *si on peut dire* (cf. l'étude de Rouanne 2016).

inattendu du moment que, comme l'indiquent les études théoriques descriptives présentées dans 2.2.5.1, *je trouve que* est compatible, de la même manière que *je pense/crois que*, avec l'expression d'une évaluation, comme dans les exemples suivants :

(106) Moi, personnellement, quand je m'y promène, **je trouve que** ce n'est pas mal. (PRES FR)

(107) Et **je trouve que** ces risques sont extrêmement inquiétants. (LM)

Néanmoins, dans nos corpus, *je trouve que* n'a pas de préférence statistique à introduire des évaluations exprimées par des adjectifs ou des adverbes subjectifs. Une explication peut être le fait que les mouvements évaluatifs liés à *je trouve que* sont très variés (en termes de marqueurs) et par conséquent aucun motif particulier se dessine, à la différence des motifs très standardisés observés pour *je pense que* et *je crois que* dans les exemples (85) à (88).

Un autre résultat qui différencie *je trouve que* de *je pense que* et de *je crois que* est le fait que *je trouve que* ne présente pas de spécificité avec des marqueurs déontiques, bien que ce verbe ne soit pas incompatible avec des jugements normatifs tels que ceux exprimés ci-dessous :

(108) **Je trouve qu'il ne faut pas** mélanger les genres. (PRES-FR)

(109) Pour l'instant, c'est une question de volonté et moi, **je trouve que** les jeunes [...] doivent apprécier ce qui s'appelle l'effort. (PRES-FR)

On constate que dans ces exemples *je trouve que* joue un rôle similaire à celui que jouent *je pense que* et *je crois que* dans les exemples (89) à (100) : le locuteur introduit des interdictions (*il ne faut pas mélanger les genres*) ou des obligations (*les jeunes doivent apprécier ce qui s'appelle l'effort*) en les prenant personnellement en charge. Les outils statistiques nous montrent toutefois que ce processus d'introduction de jugements normatifs est beaucoup plus fréquent et systématisé avec *je pense que* et *je crois que* qu'avec *je trouve que*.

Les analyses que nous proposons ouvrent de nouvelles perspectives pour différencier *je trouve que* de *je pense/crois que* : en plus des distinctions mises en lumière par Ducrot (1975) et Gosselin (2015), nous pouvons avancer que *je trouve que* ne montre pas de préférence particulière quant à la sélection de marqueurs évaluatifs et n'introduit pas de fréquemment des jugements déontiques. Ces tendances devraient être vérifiées dans d'autres corpus afin de déterminer si elles sont dues au sémantisme du verbe ou si elles sont spécifiques aux corpus étudiés.

4.3.2.3 *Je considère que et je dis que*

Dans PRES-FR, *je considère que* et *je dis que* ont des caractéristiques semblables à celles mises en exergue pour *je pense que* et *je crois que*, car nous observons une variété (plus ou moins riche selon la forme) de marqueurs dans les quatre classes (cf. Figure 14 et Figure 15). Cette sensibilité aux CS ne se retrouve néanmoins pas dans le corpus LM, ce qui nous empêche de parvenir à des conclusions claires en ce qui concerne les propriétés de ces deux verbes d'indication de point de vue.

Toutefois, si nous nous focalisons sur la classe *Modalité*, une régularité se dessine dans les deux corpus, à savoir l'attraction statistiquement significative de *je considère que* avec le verbe *devoir* et celle de *je dis que* avec *falloir*. Nous pouvons alors avancer que, comme *je pense que* et *je crois que*, *je considère que* et *je dis que* s'insèrent tout aussi bien dans des mouvements argumentatifs introduisant des demandes de faire, mais se spécialisent chacun sur un verbe précis :

- (110) C'est une décision suffisamment grave pour que je demande au Premier ministre, au ministre de la Défense, et au ministre des Affaires étrangères d'informer les commissions parlementaires. Parce que **je considère que** nous devons frapper, et frapper fort, au moment où Daech, le groupe de l'Etat islamique, commence à fléchir et à se retirer. (PRES-FR)
- (111) Comprenez-moi bien, moi aussi **je considère qu'on** doit mieux sécuriser la frontière. (LM)
- (112) De l'argent il en faut, c'est un moyen d'échange nécessaire dans une société moderne. **Je dis qu'il** faut absolument que cet argent soit orienté vers les investissements qui servent notre pays, parce qu'ils servent notre économie et qu'il faut prendre tous les moyens qui conviendront pour éviter l'argent spéculatif. (PRES-FR)
- (113) « Ici, si tu n'as pas une barque et un cheval, tu n'es pas un homme, affirme le vieux pêcheur Vladimir Nistor. Mais les chevaux sauvages, qu'ils aillent au diable ! Moi, **je dis qu'il** faut tous les tuer, ils bouffent nos forêts. » (LM)

Dans l'exemple (110), le locuteur appelle la mise en place d'une action militaire (*nous devons frapper*) et assume totalement la responsabilité de cette prise de position grâce au verbe d'indication de point de vue (*je considère que*) ; dans (111), il se montre en accord avec son interlocuteur (*moi aussi je considère que*) sur la nécessité (*on doit*) de mieux sécuriser la frontière. Dans les extraits (112) et (113), il demande respectivement que l'argent soit orienté vers des investissements productifs et que les chevaux sauvages soient abattus au moyen du verbe *falloir*, qui exprime une nécessité impersonnelle et objective : *je dis que* vient alors personnaliser ce contenu en faisant comprendre qu'il le prend en charge.

Comme *je pense/crois que*, ces deux verbes d'indication de point de vue viennent donc introduire, de façon prototypique, des jugements normatifs. Mais à la différence de *je pense/crois que*, il nous semble que lorsque *je considère que* et *je dis que* sont employés, comme dans les exemples (110) à (113), une lecture modale soit difficilement acceptable. En effet, lorsque le locuteur utilise des formes comme *je considère que* ou *je dis que*, c'est λ qui est mis en scène : il décrit son processus de réflexion cognitive ou sa parole. Cette manifestation de λ rend plus difficile la modification rétroactive de l'intensité de la demande de faire. Observons :

(114) **Je pense/crois que** nous devons frapper, *mais peut-être que je me trompe.*

(115) **?Je considère/dis que** nous devons frapper, *mais peut-être que je me trompe.*

La suite *mais peut-être que je me trompe* permet de revenir sur la demande de faire, de la remodeler : si l'énoncé (114) ne pose pas de problème (*je pense/crois que* peuvent appeler une interprétation modale qui est compatible avec *mais peut-être que je me trompe*), l'exemple (115) est plus difficilement acceptable car la suite qui estompe l'ordre se heurte au caractère réfléchi de l'ordre lui-même et entre donc en contradiction avec l'intensité du but illocutoire de l'énoncé que *je considère/dis que* aident à construire. En conclusion, quand *je considère que* et *je dis que* sont associés à des verbes déontiques, leur fonction est celle de préparer un jugement normatif (comme *je pense que* et *je crois que*), mais dans le but d'insister sur la demande de faire en augmentant la force du but illocutoire de l'énoncé. Nous sommes donc en présence d'une lecture attributive, qui est par ailleurs favorisée soit par l'utilisation du pronom *moi* – (111) et (113) – soit par l'exploitation du potentiel argumentatif de *parce que* ainsi que de la reprise du verbe *frapper* modifié par l'adverbe *fort* (110), soit par le

renforcement du verbe *falloir* par l'adverbe *absolument* (112) : dans tous ces exemples, le but du locuteur n'est pas de présenter *p*, mais de se mettre en scène en tant qu'instance à l'origine de l'obligation.

4.3.2.4 *Je suis sûr·e que et je suis convaincu·e que*

Bien qu'à travers *je suis sûr·e que* et *je suis convaincu·e que* le locuteur prenne la décision de se mettre explicitement en jeu dans son discours et de présenter *p* comme ayant un degré de validation élevé, ces deux verbes d'indication de point de vue se distinguent nettement des autres verbes du paradigme, car ils n'atteignent systématiquement pas le seuil de spécificité avec la macro-catégorie des CS correspondante (cf. les échelles de mise en scène énonciative dans la section 4.1).

Ce qui explique ces tendances est le fait que, dans nos corpus, *je suis sûr·e que* et *je suis convaincu·e que* ne sont pas mobilisés pour véhiculer une évaluation personnelle, mais plutôt pour livrer des suppositions concernant un état de choses à venir. Cette hypothèse naît de l'analyse des cooccurrents morphosyntaxiques de *je suis sûr·e que* et *je suis convaincu·e que*, qui met en évidence le fait que la seule catégorie verbale à laquelle ils sont statistiquement associés est celle des verbes au futur de l'indicatif⁴⁷. Quelques exemples tirés des deux corpus illustrent ce fonctionnement :

- (116) Que vienne la paix des braves, et **je suis sûr que** les haines iront en s'effaçant. (PRES-FR)
- (117) A l'heure qu'il est, on en est au rythme de 400000 logements par an, et **je suis convaincu qu'on va arriver** dans un, deux ou 3 ans, à 500000 logements par an. (PRES-FR)
- (118) J'ai averti Silvio Berlusconi de la situation et **je suis sûr qu'il soutiendra** notre intérêt et celui du pays. (LM)

⁴⁷ Ce comportement a été observé en prenant en compte un empan droit de 10 items. Dans PRES-FR, la catégorie des verbes au futur de l'indicatif présente un indice Lafon de 19 avec *je suis sûr·e que* et de 5 avec *je suis convaincu·e que*. Dans LM, les résultats sont confirmés pour *je suis sûr·e que* (indice Lafon de 3), bien que pour *je suis convaincu·e que* nous observons également une attraction à la catégorie du conditionnel présent (pour les deux l'indice Lafon est de 2).

- (119) Etant donné les difficultés énormes qu'impliquent les différentes alternatives, **je suis convaincu que** la Grèce essaiera d'endurer l'épreuve aussi longtemps que possible. (LM)

Dans tous ces exemples, *je suis sûr que* et *je suis convaincu que* sont employés pour traiter d'événements futurs et traduisent une certaine certitude du locuteur par rapport à ces situations, qui peut être interprétée comme une forme de souhait : que la haine disparaisse (116), que le nombre de logements augmente (117), que Berlusconi soutienne l'intérêt du pays (118), que la Grèce sorte positivement des difficultés économiques qu'elle endure (119).

4.3.2.5 *Il me semble que*

Il me semble que a un comportement très instable selon les corpus. Dans LM, il n'atteint pas le seuil de spécificité avec la macro-catégorie de CS correspondante (comme *je suis sûr-e que* et *je suis convaincu-e que*), mais dans PRES-FR oui, et cette spécificité de la macro-catégorie est due à un seul et unique marqueur : le verbe *devoir*. L'analyse des cooccurrents permet de constater que tous les emplois du verbe *devoir* en cooccurrence avec *il me semble que* dans PRES-FR sont des emplois déontiques, comme dans les exemples suivants :

- (120) D'abord l'idée que j'ai de ce que devraient être les rapports entre l'Etat et la presse : **il me semble que** ces rapports devraient être aussi peu nombreux que possible. (PRES-FR)
- (121) Je suis très prudent sur tous les rapports d'experts et moi **il me semble que** l'avenir doit en tous les cas rendre très attentif à ne pas isoler la France des grands courants d'échanges. (PRES-FR)
- (122) Je vous citerai deux sujets à propos desquels **il me semble que** des initiatives doivent être prises. (PRES-FR)

Le motif linguistique reliant *il me semble que* à *devoir* dans PRES-FR atteste que ce verbe d'indication de point de vue permet, tout comme *je pense que* et *je crois que*, une prise de distance du locuteur par rapport aux jugements normatifs qu'il avance dans *p* (l'interprétation est systématiquement modale). Toutefois, contrairement à *je pense que* et *je crois que*, avec *il me semble que* ces jugements sont prototypiquement exprimés par le verbe *devoir* uniquement. Il faut néanmoins tenir compte du fait que ce comportement est relatif au corpus PRES-FR

uniquement et que pour confirmer ces hypothèses, il faudrait approfondir cette question en prenant en compte d'autres corpus.

4.3.3 Synthèse du niveau micro

Le niveau micro a permis d'étudier de manière détaillée le profil combinatoire des verbes d'indication de point de vue, à savoir les environnements lexicaux qui les caractérisent de manière statistiquement significative. *Je pense que* et *je crois que* sont les formes qui ont le plus de CS dans leur entourage proche : elles ont tendance à apparaître avec d'autres marqueurs appelant une posture investie du locuteur et sont très fréquemment utilisées pour préparer des jugements normatifs. *Je trouve que* se distingue de ces deux formes car les cooccurrents des classes *Evaluation* et *Modalité* sont moins nombreux, voire absents. *Je considère que* et *je dis que* se lient de façon prototypique aux modalités déontiques exprimées respectivement par *devoir* et *falloir*. Enfin, *je suis sûr-e que* et *je suis convaincu-e que* sont les verbes d'indication de point de vue qui présentent le moins de CS et *il me semble que* a un comportement très instable selon le corpus dans lequel il est utilisé.

5. Présentation des résultats de l'analyse cooccurentielle : les verbes d'indication de point de vue italiens

Les études sur les verbes d'indication de point de vue italiens demeurent encore très limitées ; celles qui existent, et que nous avons présentées dans la section 2.3, s'appuient sur les travaux pionniers du domaine anglophone pour l'analyse des formes *penso* et *credo* (Venier 1991, Cignetti 2009). D'autres études explorent des classes plus larges de verbes – parmi lesquels nous retrouvons plusieurs des verbes d'indication de point de vue de notre paradigme – d'un point de vue sémantique (Strik Lievers 2012) ou sémantico-pragmatique (Lo Baido 2024).

Cette étude permet d'apporter des éléments de description supplémentaires à ces verbes, notamment à un niveau sémantico-pragmatique, tout en élargissant le paradigme des formes prises en compte. La démarche d'investigation pour les verbes italiens est identique à celle utilisée pour le français : un profil combinatoire des cooccurents statistiquement significatifs est dressé pour chaque verbe. Cette analyse cooccurentielle se décline en trois niveaux : le niveau macro présente les échelles de mise en scène énonciative, le niveau intermédiaire examine les différentes classes sémantiques qui caractérisent les verbes à travers des graphiques à bâtons, et enfin le niveau micro se focalise de manière plus spécifique sur le lexique, en examinant les cooccurents spécifiques sélectionnées pour chaque verbe.

5.1 Le niveau macro

Le niveau macro permet d'évaluer la sensibilité des verbes d'indication de point de vue italiens à la macro-catégorie des cooccurents spécifiques sélectionnés (CS). Les verbes sont présentés dans une échelle de mise en scène énonciative et leur force d'association avec la macro-catégorie correspondante est indiquée par l'indice du *log-likelihood*. Rappelons qu'une association statistiquement significative est signalée par un *log-likelihood* égal ou supérieur à 10,83 ; en dessous de cette valeur, l'association n'est donc pas statistiquement démontrée. Nous présentons et commentons tout d'abord les résultats pour le corpus PRES-IT.

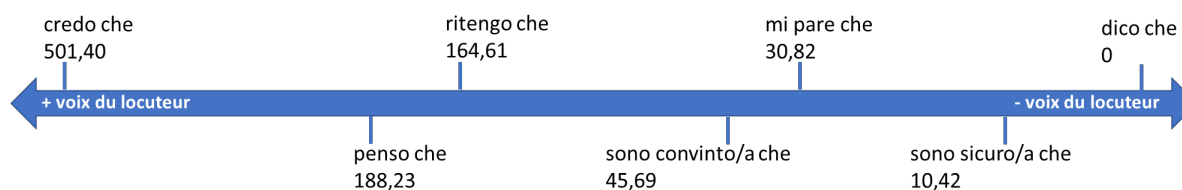


Figure 17 : échelle de mise en scène énonciative des verbes d'indication de point de vue dans PRES-IT.

L'échelle de mise en scène énonciative pour le corpus des allocutions des présidents italiens présente, au premier rang, *credo che* avec un *log-likelihood* de 501,40. Il s'agit d'une valeur extrêmement élevée du moment que le seuil de spécificité de cet indice est fixé à 10,83. Au deuxième rang, nous trouvons *penso che* (*log-likelihood* de 188,23) et au troisième *ritengo che* (*log-likelihood* de 164,61). Au milieu de l'échelle, apparaissent *sono convinto/a che* (*log-likelihood* de 45,69) et *mi pare che* (*log-likelihood* de 30,82). Les deux verbes qui se trouvent au pôle négatif de l'échelle (*sono sicuro/a che*, *dico che*) ne sont pas en situation de spécificité significative avec la macro-catégorie des CS. Passons maintenant à l'échelle pour le corpus REP :

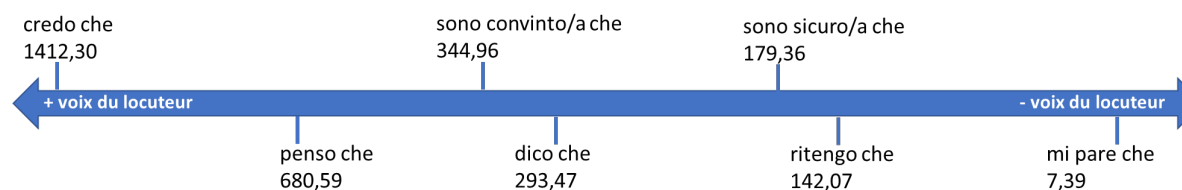


Figure 16 : échelle de mise en scène énonciative des verbes d'indication de point de vue dans REP.

Dans le corpus de presse nationale (REP), nous retrouvons les mêmes verbes aux deux premiers rangs : *credo che* (1412,30) et *penso che* (680,59). Au centre de l'échelle apparaissent *sono convinto/a che* (344,96), *dico che* (293,47), *sono sicuro/a che* (179,36) et *ritengo che* (142,07). Nous remarquons que les indices de *log-likelihood* sont beaucoup plus élevés dans ce corpus que dans PRES-IT. Le seul verbe qui ne présente pas d'attraction avec la macro-catégorie des CS est *mi pare che* (7,39).

Les différences majeures entre les deux corpus italiens concernent *dico che*, *sono sicuro/a che* et *mi pare che*. Les deux premiers ne sont pas spécifiques à la macro-catégorie des CS dans PRES-IT, mais le sont dans REP. La situation s'inverse pour *mi pare che*, qui est spécifique à la macro-catégorie des CS dans PRES-IT mais ne l'est pas dans REP. Malgré ces différences, nous observons une forte sensibilité générale des verbes d'indication de point de vue aux marqueurs de mise en scène énonciative, dans les deux corpus. *Credo che*, en particulier, se

distingue au sein du paradigme pour ses indices extrêmement importants (501,40 dans PRES-IT et 1412,30 dans REP).

5.1.1 Interprétation des échelles de mise en scène énonciative

Les échelles de mise en scène énonciative permettent de montrer que les verbes d'indication de point de vue italiens ne réagissent pas de manière homogène face à la macro-catégorie des CS. *Credo che, penso che, sono convinto/a che* et *ritengo che* affichent une tendance statistique marquée à apparaître dans des contextes où le locuteur marque explicitement sa présence, bien plus que d'autres verbes du même paradigme. En effet, leur comportement est stable dans les deux corpus et les exemples suivants illustrent leur affinité systématique avec des marqueurs qui mettent en scène explicitement la voix du locuteur⁴⁸ :

- (123) Francamente, **credo che** sia stato molto più utile rappresentare il quadro che abbiamo ascoltato da lui e dal Professore Cingolani. (PRES-IT)

Franchement, je crois qu'il a été beaucoup plus utile de représenter le cadre que nous avons entendu de lui et du professeur Cingolani.

- (124) Dejan Stankovic sottolinea la correttezza delle squadre in campo: **“Penso che** abbiamo dato una grande mano all'arbitro che è stato bravo”. (REP)

Dejan Stankovic a souligné l'équité des équipes sur le terrain : « Je pense que nous avons donné un grand coup de main à l'arbitre qui a été bon ».

- (125) **Sono convinto che** sotto l'autorevole guida del Presidente Lula, il Brasile continuerà ad essere punto di riferimento imprescindibile per una sempre più proficua collaborazione fra Europa e America Latina, favorendo una più stretta integrazione all'interno anche del sub-continente americano e una maggiore interazione fra queste due grandi aree. (PRES-IT)

Je suis convaincu que, sous la présidence experte de Monsieur Lula, le Brésil continuera à être un point de référence indispensable pour une collaboration de plus

⁴⁸ Rappelons que sur la base des critères méthodologiques établis dans la section 3.4.2.2, seuls les marqueurs soulignés ont pu être pris en compte dans l'analyse.

en plus fructueuse entre l'Europe et l'Amérique latine, en favorisant une intégration plus étroite également au sein du sous-continent américain et une plus grande interaction entre ces deux grandes régions.

- (126) **Ritengo che** la decisione della maggioranza sia sbagliata e apra un notevole conflitto d'interessi mettendo alla guida della commissione ambiente un esponente politico che nella sua attività economica privata ha interessi nel campo dell'urbanistica spesso in aperto contrasto con la difesa dell'ambiente. (REP)

Je considère que la décision de la majorité est erronée et ouvre un conflit d'intérêts considérable en plaçant à la tête de la commission de l'environnement un homme politique dont les intérêts commerciaux privés se situent dans le domaine de l'urbanisme, souvent en conflit ouvert avec la protection de l'environnement.

Dans tous ces exemples, le locuteur se met en jeu non seulement par l'utilisation du verbe d'indication de point de vue en question, mais aussi par le choix d'autres marqueurs : *francamente, molto più utile* (123), *una grande mano, bravo* (124), *autorevole, imprescindibile, sempre più proficua, una più stretta, maggiore, grandi* (125), *sbagliata, notevole* (126). En l'absence d'outils quantitatifs, il n'aurait pas été possible de prévoir que certains verbes apparaissent de façon prototypique dans ce type de contextes, contrairement à d'autres. En effet, même les verbes qui ne présentent pas d'association statistiquement positive avec la macro-catégorie des CS (*sono sicuro/a che* et *dico che* dans PRES-IT, et *mi pare che* dans REP) ne sont pas pour autant incompatibles avec des marqueurs de mise en scène de la voix du locuteur. Les exemples suivants, tirés de nos corpus, en témoignent :

- (127) Per il momento siamo ai colloqui iniziali, ma sono sicuro che la sua conclusione potrà rappresentare un contributo rilevante per la promozione di nuovo sviluppo, di avanzamento tecnologico e giustizia sociale su entrambe le sponde dell'Atlantico. (PRES-IT)

Pour l'instant, nous en sommes aux premiers pourparlers, mais je suis sûr que sa conclusion sera une contribution majeure à la promotion d'un nouveau développement, du progrès technologique et de la justice sociale de part et d'autre de l'Atlantique.

- (128) Piuttosto che dare assicurazioni, io dico che le misure prese sono state serie, per quello che riguarda l'Italia. (PRES-IT)

Plutôt que de rassurer, je dis que les mesures prises ont été sérieuses, en ce qui concerne l'Italie.

(129) “**Mi pare che** la situazione sia arrivata a un punto insostenibile, credo che la Polverini stessa abbia fatto un gesto che va comunque sottolineato” riconosce il segretario del Pd, Pier Luigi Bersani. (REP)

« Il me semble que la situation a atteint un point intenable, je crois que Madame Polverini elle-même a fait un geste qui se doit d’être souligné », reconnaît le secrétaire du PD, Pier Luigi Bersani.

Dans les exemples ci-dessus, on retrouve des adjectifs comme *rilevante* (127), *serie* (128) et *insostenibile* (129) – dans ce dernier exemple, soulignons également la présence de *credo che* qui vient s’ajouter à la suite du discours. Dans les exemples (127) et (128) le connecteur *ma* et le pronom *io* rendent également manifeste la voix du locuteur. En ce sens, rien ne distingue les contextes des exemples (127) à (129) des contextes des exemples (123) à (126). Pourtant, les outils statistiques révèlent que *sono sicuro/a che*, *dico che* et *mi pare che* ont un comportement instable dans les deux corpus, ce qui permet d’observer qu’il n’y a pas de standardisation similaire à celle de *credo che*, *penso che*, *sono convinto/a che* et *ritengo che*, qui se particularisent, à l’intérieur du paradigme, par une présence beaucoup plus systématique dans des contextes de mise en scène énonciative.

5.1.2 Synthèse du niveau macro

Les échelles de mise en scène énonciative montrent que les verbes d’indication de point de vue italiens se différencient entre eux par une sensibilité différente aux contextes où le locuteur se met scène. Les tendances sont stables dans les deux corpus pour *credo che*, *penso che*, *ritengo che* et *sono convinto/a che*, lesquels montrent une tendance statistique forte à apparaître dans ce type de contextes. Les formes *mi pare che*, *sono sicuro/a che* et *dico che* ont un comportement plus instable, qui varie d’un corpus à l’autre. Le niveau intermédiaire permettra d’approfondir ces similarités et différences en investiguant le rôle des classes sémantiques dans la construction de la macro-catégorie des CS.

5.2 Le niveau intermédiaire

Afin d'identifier plus précisément le type sémantique de marqueurs qui composent chaque macro-catégorie, nous présentons à présent les résultats pour le niveau intermédiaire. Nous rappelons ci-dessous les quatre classes sémantiques qui ont été définies pour le français et qui s'appliquent également pour l'italien :

1. La classe ***Construction argumentative*** comprend les CS qui permettent d'établir des liens argumentatifs entre unités linguistiques du discours ;
2. La classe ***Prise en charge*** comprend les CS qui sont exploités par le locuteur pour se déclarer en tant que source responsable du contenu mis en scène ;
3. La classe ***Modalité*** comprend les CS qui expriment une modalité du possible ou du nécessaire ;
4. La classe ***Evaluation*** comprend les CS utilisés pour véhiculer une appréciation ou une émotion par rapport à un contenu donné.

Le Tableau 10 dans les Annexes illustre, pour chaque marqueur rencontré, la classe dans laquelle il a été inséré. Comme pour le français, les marqueurs des classes *Construction argumentative* et *Prise en charge* concernent le contexte gauche des verbes d'indication de point de vue et ceux des classes *Modalité* et *Evaluation* le contexte droit ; la force d'attraction entre chaque verbe et les classes sémantiques correspondantes est indiquée par l'indice de *log-likelihood*. Le graphique suivant présente les résultats pour le corpus PRES-IT :

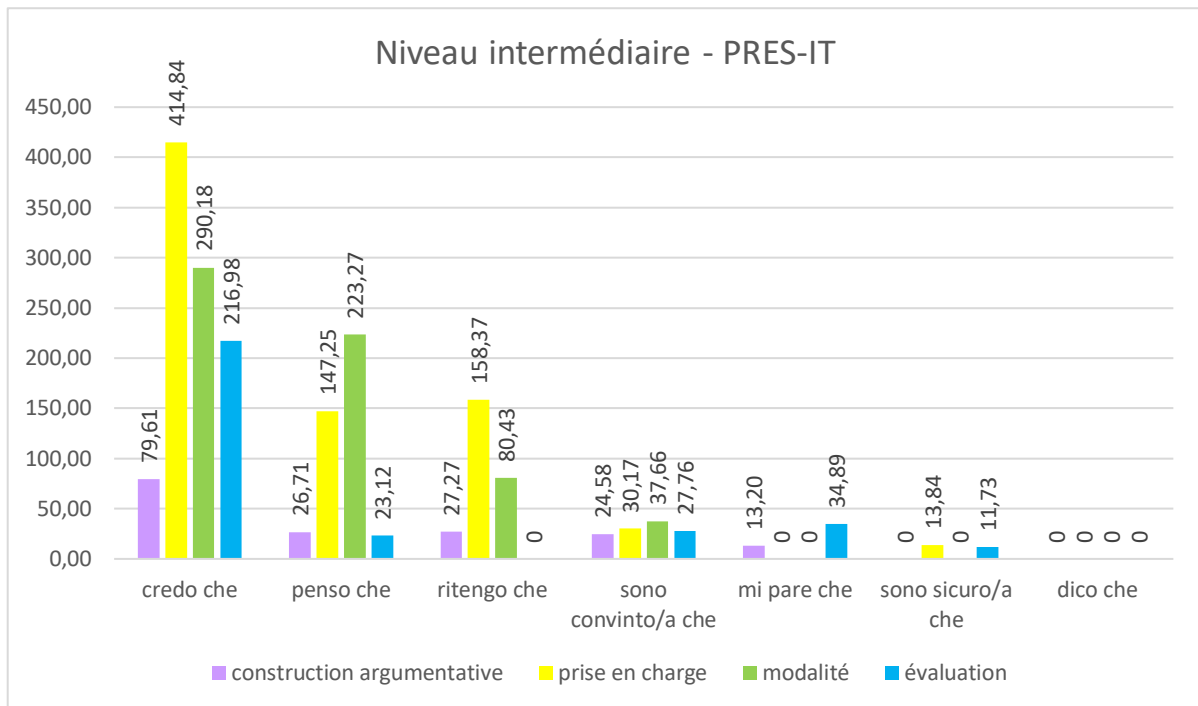


Figure 18 : classes sémantiques spécifiques aux verbes d'indication de point de vue dans PRES-IT.

Nous observons que les deux verbes d'indication de point de vue qui apparaissent au pôle positif de l'échelle de mise en scène énonciative du niveau macro (*credo che* et *penso che*) sont attirés par les quatre classes sémantiques ; il en va de même pour *sono convinto/a che*. *Ritengo che* est sensible à toutes les classes sauf *Evaluation* et *mi pare che* aux classes *Construction argumentative* et *Evaluation* uniquement. Les deux verbes qui, dans le niveau macro, ne montrent pas d'attraction statistiquement significative avec la macro-catégorie des CS sont respectivement spécifiques aux classes *Prise en charge* et *Evaluation* (*sono sicuro che*) ou à aucune classe (*dico che*). Passons à présent aux résultats pour le corpus de presse nationale (REP), afin de vérifier si les mêmes tendances se dégagent.

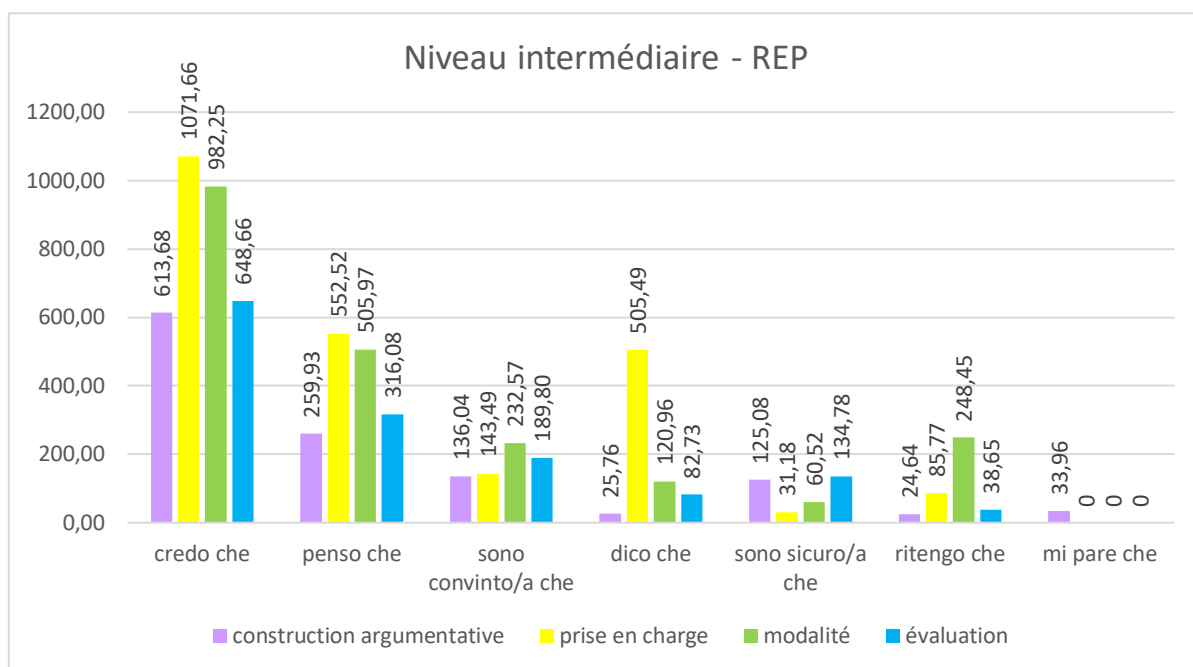


Figure 19 : classes sémantiques spécifiques aux verbes d'indication de point de vue dans REP.

Dans ce corpus, tous les verbes d'indication de point de vue à l'exception de *mi pare che* (le seul verbe à ne pas être spécifique à la macro-catégorie des CS) sont attirés par les quatre classes sémantiques. Les indices des classes sont d'ailleurs très forts : dans 61% des cas ils dépassent la valeur de 100. Les points communs observés entre les deux corpus se concentrent autour des trois verbes d'indication de point de vue *credo che*, *penso che* et *sono convinto/a che*, qui sont chacun associé aux quatre classes sémantiques. Les divergences majeures se manifestent principalement pour *sono sicuro/a che* et *dico che*, qui, dans le corpus REP, attirent les quatre classes, contrairement à ce qui se passe dans PRES-IT (le premier attire deux classes et le second aucune).

5.2.1 Interprétation des graphiques du niveau intermédiaire

Trois points principaux ressortent de l'analyse des graphiques du niveau intermédiaire :

- Dans les deux corpus, **les verbes qui occupent les rangs supérieurs des deux échelles du niveau macro (*credo che*, *penso che*, *sono convinto/a che*) sont attirés par les quatre classes sémantiques.** Les classes *Prise en charge* et *Modalité* ont des indices très élevés, qui dépassent de façon constante (exception faite pour la classe *Evaluation* de *sono convinto/a che* dans REP) les indices des deux autres catégories : cela peut être dû à une extrême richesse des marqueurs composant ces deux classes, ou alors à

des marqueurs spécifiques pouvant avoir des indices très élevés. L'analyse encore assez générale de ce niveau ne permet pas de comprendre ces tendances, mais elles seront approfondies et expliquées dans le niveau micro (cf. section 5.3).

- Dans les deux corpus, **les verbes qui occupent les rangs inférieurs des échelles**, à savoir ceux qui ne sont pas spécifiques à la macro-catégorie (*sono sicuro/a che* et *dico che* dans PRES-IT et *mi pare che* dans REP) et ceux qui présentent des indices relativement bas (*mi pare che* dans PRES-IT), **effectuent des sélections précises des classes**, ou n'en sélectionnent aucune.
- Dans le corpus REP, lorsqu'un verbe est spécifique à la macro-catégorie des CS, toutes les classes sémantiques sont systématiquement représentées. Cela pose la question de savoir à quel point l'usage des verbes est influencé par le type de texte dans lequel il apparaît. Ces questions seront approfondies dans le niveau micro, notamment pour les verbes qui présentent des comportements opposés d'un corpus à l'autre : *sono sicuro/a che* et *dico che*, qui ne sont pas spécifiques à la macro-catégorie des CS dans PRES-FR mais qui le sont dans REP, où ils sont sensibles aux quatre classes sémantiques.

5.2.2 Synthèse du niveau intermédiaire

L'analyse du niveau intermédiaire a permis de déterminer les classes sémantiques principalement responsables des indices des macro-catégories dans les échelles de mise en scène énonciative pour les verbes d'indication de point de vue italiens. Les verbes situés aux rangs supérieurs des deux échelles (*credo che*, *penso che*, *sono convinto/a che*) montrent une sensibilité à toutes les classes sémantiques – avec des indices particulièrement élevés pour les classes *Prise en charge* et *Modalité* – tandis que ceux aux rangs inférieurs tendent à sélectionner précisément une ou plusieurs classes. Le comportement divergent entre verbes d'indication de point de vue dans les deux corpus soulève des questions sur l'influence du type de texte sur l'utilisation de ces formes, questions qui seront approfondie au niveau micro.

5.3 Le niveau micro

Le troisième niveau d'analyse offre une vision détaillée sur les marqueurs qui cooccurrent avec les verbes d'indication de point de vue. En se concentrant spécifiquement sur le lexique, ce niveau permet en effet de clarifier la composition des classes sémantiques. Comme pour le français, les résultats sont présentés pour chaque corpus dans des tableaux différents. Les

verbes d'indication de point de vue (au centre) sont classés selon leur position sur l'échelle de mise en scène énonciative du niveau macro (les plus énonciatifs en haut, les moins énonciatifs en bas). Les deux premières colonnes (*Construction argumentative* et *Prise en charge*) répertorient les marqueurs cooccurents avec chaque verbe dans l'empan gauche de 4 items. Les colonnes de droite (*Modalité* et *Évaluation*) présentent les marqueurs en cooccurrence avec les verbes dans l'empan droit de 10 items. Chaque marqueur est reporté sous forme de lemme, accompagné de la catégorie morphosyntaxique correspondante et de l'indice de cooccurrence de Lafon (1981), qui mesure la force d'attraction avec le verbe correspondant (rappelons que, pour qu'il y ait spécificité, il faut un indice égal ou supérieur à la valeur-seuil de 3). Nous commençons par présenter les résultats du corpus des allocutions présidentielles (PRES-IT).

CONSTRUCTION ARGUMENTATIVE		PRISE EN CHARGE			MODALITE		EVALUATION	
marqueur	indice de spécificité	marqueur	indice de spécificité		marqueur	indice de spécificité	marqueur	indice de spécificité
ADV_però	11,33	PRO:pers_io	87,18	CREDO CHE	VER_potere	30,16	ADV_veramente	17,01
CON_ma	10,79	ADV_francamente	4,78		VER_dovere	16,48	ADV_molto	10,22
					VER_bisognare	8,35	ADJ_importante	9,03
					NOM_bisogno	4,04	ADV_davvero	7,44
							ADJ_justo	4,70
							ADV_seramente	4,51
CON_perché	4,37	PRO:pers_io	33,05	PENSO CHE			ADV_ragionevolmente	3,92
CON_ma	3,40				VER_potere	14,99	ADJ_buono	5,80
					VER_dovere	12,92		
					VER_bisognare	5,08		
CON_perché	6,72	PRO:pers_io	35,43	RITENGO CHE	ADV_forse	4,84		
CON_ma	3,75	PRO:pers_io	7,36		VER_dovere	5,66		
CON_perché	3,22			SONO CONVINTO/A CHE	VER_potere	6,44	ADJ_grande	5,30
CON_perché	3,55						ADV_veramente	3,04
				MI PARE CHE			ADV_molto	5,82
							ADV_già	3,31
		PRO:pers_io	3,64	SONO SICURO/A CHE			ADJ_duraturo	3,11
				DICO CHE				

Figure 20 : cooccurents spécifiques retenus pour chaque verbe d'indication de point de vue dans PRES-IT.

La composition des classes sémantiques pour les verbes d'indication de point de vue dans PRES-IT présente des variations significatives. Certains verbes, notamment *credo che* et *penso che*, se distinguent par une richesse de marqueurs bien plus grande que les autres. *Credo che* inclut des marqueurs appartenant aux quatre classes : les conjonctions *però* (mais, toutefois) et *ma* (mais) pour la classe *Construction argumentative*, le pronom personnel *io* (moi) et l'adverbe *francamente* (franchement) pour la classe *Prise en charge*, les verbes modaux *potere* (pouvoir), *dovere* (devoir), *bisognare* (falloir) ainsi que le nom *bisogno* (besoin)

pour la classe *Modalité*, et les adverbes *veramente* (vraiment), *molto* (beaucoup), *davvero* (vraiment, réellement, très), *seriamente* (sérieusement), *ragionevolmente* (raisonnablement), *ben* (bien, très), *fortemente* (fortement) et les adjectifs *importante* (important) et *giusto* (juste) pour la classe *Évaluation*. *Penso che* est également sensible aux quatre classes sémantiques et partage la majorité de ses marqueurs avec *credo che*. Ses cooccurrents spécifiques incluent *perché* (parce que) et *ma* (mais) pour la *Construction argumentative*, *io* (moi) pour la *Prise en charge*, *potere* (pouvoir), *dovere* (devoir), *bisognare* (falloir), *forse* (peut-être) pour la *Modalité*, et *buono* (bon) pour l'*Évaluation*. *Sono convinto/a che* est le troisième verbe à être sensible aux quatre classes et partage en grande partie les mêmes marqueurs que *credo che* et *penso che* (nous retrouvons en effet les marqueurs *ma*, *perché*, *io*, *potere*, *veramente*).

Les autres verbes (*ritengo che*, *mi pare che*, *sono sicuro/a che*) possèdent moins de marqueurs spécifiquement associés (parmi ceux-ci apparaissent néanmoins encore une fois *perché*, *io*, *dovere*), tandis que *dico che* n'en a aucun. Examinons à présent les résultats pour le corpus REP.

CONSTRUCTION ARGUMENTATIVE		PRISE EN CHARGE			MODALITE		EVALUATION		
marqueur	indice de spécificité	marqueur	indice de spécificité		marqueur	indice de spécificité	marqueur	indice de spécificité	
CON_ma	124,16	PRO:pers_io	212,92	CREDO CHE	VER_potere	102,91	ADJ_justo	28,45	
CON_perché	11,83	ADV_personalmente	11,09		VER_dovere	74,67	ADV_molto	27,97	
CON_però	4,83	ADV_francamente	5,02		NOM_bisogno	6,70	ADV_bene	24,22	
					VER_bisognare	6,25	ADJ_importante	12,62	
					ADJ_necessario	5,53	ADJ_migliore	11,85	
					NOM_possibilità	5,14	ADJ_buono	8,46	
					VER_occorrere	4,75	ADVdavvero	7,89	
					ADV_forse	3,26	ADV_piu	7,82	
					NOM_dovere	3,04	ADJ_ottimo	7,11	
							ADJ_buon	5,80	
							ADJ_certo	5,74	
							ADV_meglio	5,19	
							ADJ_grande	4,75	
							ADJ_accettabile	4,12	
							ADJ_difficile	4,02	
							ADJ_opportuno	3,89	
							ADJ_fondamentale	3,86	
							ADJ_sufficiente	3,79	
							ADV_sicuramente	3,76	
							ADJ_deleterio	3,51	
							ADJ_normale	3,43	
							ADJ_simile	3,30	
							ADJ_competitivo	3,26	
							ADJ_tanto	3,04	
							ADJ_antipatico	3,04	
							ADJ_maturo	3,02	
CON_ma	54,30	PRO:pers_io	116,54		PENSO CHE	VER_potere	46,71	ADV_bene	12,99
CON_perché	4,01	ADV_onestamente	5,96			VER_dovere	44,14	ADV_piu	11,86
CON_però	3,07					VER_bisognare	8,56	ADJ_importante	11,22
						NOM_possibilità	5,42	ADJ_justo	9,73
						ADJ_doveroso	3,44	ADJ_buon	7,87
								ADJ_buono	7,12
								ADJ_migliore	5,58
								ADVdavvero	4,05
								ADV_meglio	3,90
								ADJ_fondamentale	3,54
							ADV_molto	3,50	
							ADJ_competitivo	3,34	
CON_ma	28,26	PRO:pers_io	28,11	SONO CONVINTO/A CHE	VER_potere	34,36	ADV_bene	10,67	
CON_perché	3,44	ADV_anzi	4,87		VER_dovere	9,38	ADJ_buon	6,24	
							ADJ_forte	5,21	
							ADV_molto	4,93	
							ADV_benissimo	4,77	
							ADV_piu	4,64	
							ADJ_grande	4,49	
							ADJ_importante	4,46	
							ADV_cosi	4,37	
							ADJ_justo	3,81	
						ADJ_competitivo	3,54		
						ADJ_meglio	3,23		
						ADJ_rigorouso	3,19		
CON_ma	6,52	PRO:pers_io	108,02	DICO CHE	VER_dovere	14,43	ADV_meglio	7,31	
		ADV_personalmente	3,65		VER_bisognare	7,96	ADV_bene	4,96	
					VER_potere	6,48	ADJ_sicuro	3,68	
							ADJ_forte	3,53	
CON_ma	28,33	PRO:pers_io	7,60	SONO SICURO/A CHE			ADJ_bello	3,01	
					VER_potere	5,95	ADV_bene	19,45	
					NOM_bisogno	3,96	ADJ_grande	11,21	
CON_perché	4,66	PRO:pers_io	19,59	RITENGO CHE			ADJ_buon	3,32	
CON_ma	3,03				VER_dovere	13,40	ADJ_opportuno	5,05	
					VER_potere	6,84	ADJ_justo	3,43	
CON_ma	8,29			MI PARE CHE	VER_occorrere	3,98	ADJ_eccessivo	3,13	

Figure 21 : cooccurrents spécifiques retenus pour chaque verbe d'indication de point de vue dans REP.

Rappelons tout d'abord que les niveaux macro et intermédiaire ont montré que, dans REP, tous les verbes, à l'exception de *mi pare che*, sont spécifiques à la macro-catégorie des CS et que toutes les classes sémantiques sont spécifiques : le niveau micro illustre bien ces tendances, en montrant la très grande variété de cooccurrents spécifiques associés aux verbes (Figure 21). En particulier, les verbes des trois premiers rangs de l'échelle de mise en scène énonciative

(*credo che, penso che, sono convinto/a che*), ont des classes très riches : 3 CS de *Construction argumentative*, 3 CS de *Prise en charge*, 9 CS de *Modalité* et 26 CS d'*Evaluation* pour *credo che* ; 3 CS de *Construction argumentative*, 2 CS de *Prise en charge*, 5 CS de *Modalité* et 15 CS d'*Evaluation* pour *penso che* ; 2 CS de *Construction argumentative*, 2 CS de *Prise en charge*, 2 CS de *Modalité* et 13 CS d'*Evaluation* pour *sono convinto/a che*. *Dico che, sono sicuro/a che* et *ritengo che* présentent également des CS appartenant aux quatre classes, bien que moins nombreux. Très remarquable est le fait que, dans ce corpus, la plupart des verbes du paradigme partagent, pour chaque classe, des marqueurs précis : *ma* pour *Construction argumentative*, *io* pour *Prise en charge*, *potere* et *dovere* pour *Modalité*, des adverbes et adjectifs axiologiques comme *giusto, bene, buono* pour *Evaluation*.

5.3.1 Points communs entre les deux corpus italiens

Après avoir présenté les résultats pour PRES-IT et REP, nous pouvons mettre en lumière plusieurs points communs entre ces deux corpus. Ces similitudes sont particulièrement intéressantes car elles indiquent que les comportements observés dans deux corpus différents sont très probablement dus à des mécanismes inhérents au sémantisme des verbes, et non au type de texte dans lequel les verbes apparaissent.

- Tout d'abord, l'étude minutieuse des cooccurrents spécifiques que permet le niveau micro nous amène à présent à mieux comprendre la très forte sensibilité que certains verbes manifestent dans le niveau intermédiaire par rapport aux classes *Prise en charge* et *Modalité* (cf. section 5.2.1). Il s'agit, rappelons-le, de *credo che* et *penso che*, mais aussi de *sono convinto/a che* et *ritengo che*, ainsi que de *dico che* (ce dernier, dans REP uniquement). La très forte spécificité de la catégorie *Modalité* s'explique par le fait que cette classe présente un grand nombre de cooccurrents, dont certains ont des indices très élevés : par exemple, les deux premiers cooccurrents de *credo che* dans PRES-IT sont *potere* et *dovere*, avec des indices de 30,16 et 16,48 respectivement ; dans REP ces indices de cooccurrence montent à 102,91 et 74,67. Pour *penso che*, nous observons les mêmes cooccurrents spécifiques en tête, avec des indices de 14,99 (*potere*) et de 12,92 (*dovere*) dans PRES-IT, ainsi que de 46,71 (*potere*) et de 44,14 (*dovere*) dans REP. Pour la classe *Prise en charge*, c'est uniquement le cooccurrent *io* qui fait monter les indices de la classe. En effet, nous observons une très forte spécificité de ce pronom personnel avec *credo che* (87,18 dans PRES-IT et 212,92 dans REP), avec *penso che* (33,05 dans PRES-IT et 116,54 dans REP), mais aussi avec

ritengo che (35,43 dans PRES-IT et 19,59 dans REP) et *sono convinto/a che* (7,36 dans PRES-IT et 28,11 dans REP). Dans REP, cet indice est également extrêmement élevé pour *dico che* (108,02). Ces observations permettent donc de préciser que les indices des classes sémantiques du niveau intermédiaire sont le fruit non seulement du nombre de CS composant la classe, mais également de la force d'attraction statistique de chacun d'entre eux avec le verbe correspondant. Ainsi, si la classe *Modalité* présente une assez large variété de CS, cela n'est pas le cas de la classe *Prise en charge*, car les très hauts indices qui la caractérisent sont dus à un seul et unique marqueur.

- *Credo che* et *penso che* sont les verbes d'indication de point de vue qui ont la plus grande variété de CS. Pour *credo che*, nous retrouvons dans les deux corpus les marqueurs suivants : *ma* et *però* pour *Construction argumentative* ; *io* et *francamente* pour *Prise en charge* ; *potere*, *dovere*, *bisognare*, *bisogno* pour *Modalité* ; *molto*, *davvero*, *bene*, *importante* et *giusto* pour *Evaluation*. Pour *penso che*, les marqueurs partagés sont *ma* et *perché* pour *Construction argumentative* ; *io* pour *Prise en charge* ; *potere*, *dovere*, *bisognare* pour *Modalité* ; *buono* pour *Evaluation*. Enfin, ces deux formes partagent entre elles, de façon systématique, la conjonction *ma*, le pronom personnel *io*, ainsi que les verbes modaux *potere*, *dovere*, *bisognare*.
- *Ritengo che* présente dans les deux corpus les cooccurrents spécifiques *perché* (*Construction argumentative*), *io* (*Prise en charge*) et *dovere* (*Modalité*).
- *Sono convinto/a che* partage les conjonctions *ma* et *perché* (*Construction argumentative*), le pronom *io* (*Prise en charge*), le verbe *potere* (*Modalité*), l'adjectif *grande* (*Evaluation*).
- *Sono sicuro/a che* partage le pronom *io* (*Prise en charge*).

Le tableau ci-dessous présente visuellement si les tendances des verbes d'indication de point de vue dans les deux corpus convergent. Rappelons que les cases vertes indiquent que le verbe partage soit les marqueurs indiqués, soit l'absence de marqueurs (\emptyset) ; les cases rouges signalent que le comportement du verbe diverge entre PRES-IT et REP, car aucun marqueur commun a été identifié (-).

<i>Construction argumentative</i>	<i>Prise en charge</i>		<i>Modalité</i>	<i>Evaluation</i>
ma però	io francamente	Credo che	potere dovere (v.) bisognare bisogno	molto davvero bene importante giusto
ma perché	io	Penso che	potere dovere (v.) bisognare	buono
perché	io	Ritengo che	dovere (v.)	-
ma perché	io	Sono convinto/a che	potere	grande
-	∅	Mi pare che	∅	-
-	io	Sono sicuro/a che	-	-
-	-	Dico che	-	-

Tableau 5 : tableau récapitulatif des CS communs aux verbes d'indication de point de vue dans les corpus italiens.

Le Tableau 5 montre que pour *credo che*, *penso che* et *sono convinto che*, le comportement est extrêmement régulier dans les deux corpus. En effet, chaque forme partage au moins un marqueur commun au moins pour chaque classe sémantique. Les plus grandes divergences – que nous avons déjà mises en lumière dans les sections 5.1 et 5.2 – concernent essentiellement *sono sicuro che* et *dico che*. Il est donc impératif d'approfondir et d'analyser les emplois de ces deux formes dans chacun des corpus pour mieux comprendre ces différences, ce qui sera abordé dans les sections suivantes.

5.3.2 Interprétation des tableaux du niveau micro

En italien, nous observons une dichotomie assez nette entre, d'une part, les verbes d'indication de point de vue très sensibles à la totalité ou à la majorité des classes sémantiques (*credo che*, *penso che*, *sono convinto/a che*, *ritengo che*) et, d'autre part, les verbes qui n'ont pas, ou très

peu, de marqueurs communs d'un corpus à l'autre (*dico che, sono sicuro/a che*). Entre les deux se situe *mi pare che*, qui ne partage pas de marqueurs communs d'un corpus à l'autre, mais qui montre une insensibilité constante à deux des classes sémantiques (*Prise en charge* et *Modalité*).

5.3.2.1 *Credo che et penso che*

Les expressions *credo che* et *penso che* ne se limitent pas à être sensibles aux quatre classes sémantiques, mais elles s'accompagnent systématiquement de marqueurs communs, observables tant dans PRES-IT que dans REP. Parmi ces marqueurs figurent le pronom *io* (relevant de la classe *Prise en charge*), le connecteur *ma* (classe *Construction argumentative*), ainsi que les verbes *potere, dovere* et *bisognare* (classe *Modalité*). Les sections suivantes analyseront les usages de *credo che* et *penso che* en interaction avec ces marqueurs, en examinant les effets sémantico-pragmatiques résultant de ces associations.

Les associations avec les CS de la classe *Construction argumentative*

Parmi les marqueurs de la classe *Construction argumentative*, le connecteur *ma* présente une attraction statistique forte avec *credo che* et *penso che* ; mis à part pour *penso che* dans PRES-IT, le connecteur *però* intervient également de façon régulière. Les exemples suivants illustrent ces tendances :

- (130) Quando si è scelta la strada del Fiscal Compact, [...] sia il Regno Unito sia la Repubblica Ceca hanno fatto la scelta di non sottoscrivere quello speciale nuovo accordo internazionale, per varie ragioni che non sta a me indicare e che noi rispettiamo. Però, io **credo che**, nell'interesse comune di tutta l'Unione Europea, passi come quelli che ho indicato possano essere compresi anche da governi e da Paesi che non fanno parte della zona Euro. (PRES-IT)

Lorsque la voie du Fiscal Compact a été choisie, [...] tant le Royaume-Uni que la République tchèque ont choisi de ne pas signer ce nouvel accord international spécial, pour diverses raisons qu'il ne m'appartient pas d'indiquer et que nous respectons. Cependant, je crois que, dans l'intérêt commun de l'ensemble de l'Union européenne, des mesures telles que celles que j'ai indiquées peuvent également être comprises par les gouvernements et les pays qui ne font pas partie de la zone euro.

- (131) Che cosa la Germania potrebbe imparare dall'Italia? Preferirei non rispondere direttamente a questa domanda perché avrebbe qualcosa di presuntuoso. **Ma penso che** ci siano buone ragioni perché la Germania, malgrado le nostre difficoltà attuali, possa guardare all'Italia con stima e fiducia. (PRES-IT)

Que pourrait apprendre l'Allemagne de l'Italie ? Je préférerais ne pas répondre directement à cette question, car cela aurait quelque chose de présomptueux. Mais je pense qu'il y a de bonnes raisons pour que l'Allemagne, malgré nos difficultés actuelles, puisse regarder l'Italie avec estime et confiance.

- (132) I giri di pattugliamento sono iniziati lunedì scorso, con una ventina di persone, gente del posto, che dalla piazza del quartiere ha fatto il giro delle creuze [...]. La titolare della Trattoria Bruxaboschi è scettica: “È da lunedì che girano, si fanno vedere, per dimostrare che il paese non è morto, **ma credo che** ci sia anche un po' di goliardia. Comunque, non so quanto durerà e preferirei che a girare fossero le gazzelle dei carabinieri”. (REP)

Les patrouilles ont commencé lundi dernier, avec une vingtaine de personnes, des locaux, qui ont fait le tour des creuze [...] depuis la place du quartier. La propriétaire de la Trattoria Bruxaboschi est sceptique : « Ils font des tours depuis lundi, ils se montrent, pour prouver que le village n'est pas mort, mais je pense qu'il y a aussi un peu de fanfaronnade. Mais je ne sais pas combien de temps cela va durer, et je préférerais que ce soit les gazelles des carabinieri qui se promènent ».

- (133) Abbiamo fatto il massimo. Quando miglioreremo la condizione fisica potremo arrivare lontano. Anche dopo l'1-1 abbiamo avuto occasioni da gol, ci è mancata un po' di cattiveria e un po' di lucidità. **Ma penso che** tutti abbiano visto che la Fiorentina è una squadra di carattere. (REP)

Nous avons fait de notre mieux. Quand nous améliorerons notre condition physique, nous pourrons aller loin. Même après le 1-1, nous avons eu des occasions de but, nous avons manqué d'un peu de méchanceté et de lucidité. Mais je pense que tout le monde a vu que la Fiorentina est une équipe de caractère.

Ma et però, comme leur pendant français *mais*, sont dotés des propriétés focalisatrice et énonciative (cf. Rossari et al. 2022) : ils permettent de hiérarchiser les contenus et de rendre manifeste la voix du locuteur qui vient prendre la responsabilité du contenu introduit par le connecteur. Lorsque *ma* et *però* sont en jeu, le locuteur concède les points de vue qui les

précédent, pour venir ensuite les minimiser en faveur du sien, qui est mis à l'avant-plan du discours. Dans les exemples (130) à (133), le locuteur présente les contenus suivants : *le Royaume Uni et la République Tchèque ont choisi de ne pas signer un accord international* (130), *il n'y aura pas de réponse directe à la question* (131), *les gens du villages sont présents pour combattre les vols* (132), *l'équipe de la Fiorentina manque de méchanceté et de lucidité* (133). Ces contenus sont concédés, mis à l'arrière-plan, car le locuteur met en relief, au moyen de *ma/però (io) penso/credo che*, son propre point de vue. Ainsi, ce que l'interlocuteur retient de ces discours, concerne les contenus introduits par le connecteur et le verbe : *les pays qui ne font pas partie de l'UE peuvent comprendre certains accords internationaux* (130), *l'Allemagne devrait prendre exemple de l'Italie* (131), *les patrouilles des gens locaux sont un peu frivoles* (132), *la Fiorentina est une équipe qui a du caractère* (133).

Rappelons enfin les réflexions de Cignetti (2009) concernant l'apparition de *credo che p* en début d'énoncé, accompagné par des connecteurs tels que *quindi* et *dunque* (cf. section 2.3.2). Selon lui, ces deux connecteurs sont des indices de la fonction argumentative de *credo che* ; en d'autres termes, ces connecteurs imposent une lecture attributive de *credo che*. Il nous semble que cette lecture attributive peut également être induite par *ma*, comme nous l'observons dans les exemples (130) à (133).

Les associations avec les CS de la classe *Prise en charge*

En ce qui concerne la classe *Prise en charge*, c'est le pronom personnel *io* qui a une tendance statistique extrêmement forte à introduire *credo che* et *penso che*. En italien, ce pronom est très souvent omis, car la notion de personne est sous-entendue dans la suffixation morphologique du verbe. Lorsqu'il est exprimé, il sert à mettre en évidence le sujet. Ainsi, l'ajout de *io* en tête d'un verbe implique une double mise en scène du locuteur, comme nous pouvons l'observer dans les exemples suivants :

- (134) Il saper guardare come nostro comune destino questa unità nazionale è un valore che va coltivato, è un patrimonio che va rafforzato e **io credo che** vada rafforzato sia suscitando nel Nord una nuova consapevolezza di quei doveri inderogabili di solidarietà che la Costituzione repubblicana ha solennemente sancito, sia facendo il Mezzogiorno la sua parte. (PRES-IT)

Savoir regarder à cette unité nationale comme notre destin commun est une valeur qu'il faut cultiver, c'est un héritage qu'il faut renforcer, et je crois qu'il faut le

renforcer à la fois en suscitant dans le Nord une conscience nouvelle de ces devoirs inaliénables de solidarité que la Constitution républicaine a solennellement sanctionnés, et en faisant jouer au Sud le rôle qui est le sien.

- (135) Il presidente vede le cose in anticipo: io credo che Boateng da trequartista faccia bene, è un processo che va fatto piano piano. Sa benissimo che conta vincere, occorre equilibrio, rispetta le mie scelte ma condivido molto di quello che ha detto. (REP)

Le président voit les choses à l'avance : moi je crois que Boateng se débrouille bien en tant que milieu offensif, c'est un processus qui doit se faire lentement. Il sait très bien que ce qui compte est de gagner, il faut de l'équilibre, il respecte mes choix mais je partage beaucoup de ce qu'il a dit.

- (136) “[...] Terzigno aspetta da anni il ristoro che le spetta; finora ha solo pagato costi sociali altissimi”. Procuratore, se lei vivesse lì, cosa farebbe? “Io penso che quella gente è esasperata. Penso che non ne può più. Penso che due enormi discariche in un piccolo paese non siano sopportabili per nessuno. A Terzigno come altrove. Questo penso. Ma penso anche un'altra cosa: occorre assolutamente dare atto alle forze di polizia che, nella situazione data, stanno operando in un modo straordinario [...]”. (REP)

« Terzigno attend depuis des années la réparation qu'elle mérite ; jusqu'à présent, elle n'a payé que des coûts sociaux très élevés ». Monsieur le Procureur, si vous viviez là, que feriez-vous ? Moi, je pense que ces gens sont exaspérés. Je pense qu'ils n'en peuvent plus. Je pense que deux énormes décharges dans une petite ville, ce n'est supportable pour personne. À Terzigno comme ailleurs. Voilà ce que je pense. Mais je pense aussi autre chose : il faut absolument rendre hommage aux forces de police qui, dans la situation donnée, travaillent de manière extraordinaire. »

- (137) Capisco che ciò sia stato possibile, lo voglio sottolineare, innanzitutto grazie ad un corretto metodo di consultazione con l'opposizione che il governo ha seguito. Io penso che sia un esempio significativo. (PRES-IT)

Je comprends que cela a été possible, je veux le souligner, avant tout grâce à une méthode appropriée de consultation de l'opposition que le gouvernement a suivie. Moi je pense que cela est un exemple significatif.

Dans ces exemples, le pronom *io*, utilisé en position de focalisation, vient marquer le point de vue personnel du locuteur en le renforçant. Dans (134), après avoir présenté sur un mode du « trouvé » (Carel 2011a : 10), à savoir à travers un ton factuel, que l'unité nationale italienne est une valeur à renforcer, le locuteur s'engage personnellement en exprimant ce que pour lui le nord et le sud de l'Italie devraient faire pour cultiver cette unité. Dans (135), le *io* met en parallèle deux points de vue, en permettant au locuteur de se montrer parfaitement en accord avec les choix du président. Dans (136), en répondant à la question *se lei vivesse lì, cosa farebbe?* le locuteur prend la responsabilité de ce qu'il affirme au moyen de *io penso che*, et préface ensuite les points de vue successifs par d'autres *penso che*, qui viennent ainsi rythmer le discours en rappelant que les idées avancées relèvent de sa vision personnelle. Cette propriété de renforcement de point de vue du locuteur peut également être illustrée par d'autres marqueurs intervenant dans le contexte discursif des verbes d'indication de point de vue, comme l'expression parenthétique *lo voglio sottolineare* dans (137).

En observant les autres marqueurs de la classe *Prise en charge* apparaissant à côté de *penso che* et *credo che* (cf. Figure 20 et Figure 21), nous remarquons que la focalisation sur le point de vue du locuteur peut se faire, en plus du pronom *io*, par d'autres moyens, notamment par des adverbes se terminant en *-mente* comme *personalmente* (personnellement), *francamente* (franchement), *onestamente* (honnêtement). Toutefois, c'est *credo che* qui a tendance à systématiser ce procédé, notamment avec l'adverbe *francamente* que nous retrouvons comme cooccurrent spécifique dans les deux corpus :

- (138) Sappiamo che qualche commentatore straniero ha addirittura eccepito che drammatizziamo troppo, perché quello che accade sarebbe considerato normale dovunque in Europa. Francamente, **credo che** non possiamo correre dietro a comparazioni di questa natura: è bene essere il più possibile precisi e puntuali, anche nel rivedere anno per anno – rispetto a fasi precedenti – l'andamento di questo fenomeno [...].
(PRES-IT)

Nous savons que certains commentateurs étrangers ont même objecté que nous dramatisons trop, car ce qui se passe serait considéré comme normal partout en Europe. Franchement, je crois que nous ne pouvons pas courir après des comparaisons de cette nature : il est bon d'être aussi précis et ponctuel que possible, même en examinant année par année - par rapport aux phases précédentes - la tendance de ce phénomène [...].

- (139) È l'opinione di Maurizio Piazza, sindaco di Beinasco, che aggiunge: “non si tratta di una scelta felice. Un profilo più discreto sarebbe stato preferibile su una vicenda così sensibile” sostiene. Per il primo cittadino di Grugliasco, Roberto Montà, invece è un falso problema. “Francamente **credo che** sull'inceneritore ci siano questioni più importanti di una cometa. E trovo pericolosa la strumentalizzazione che si sta facendo della vicenda: serve solo a spaventare la gente”. (REP)

C'est l'avis de Maurizio Piazza, maire de Beinasco, qui ajoute : « ce n'est pas un choix heureux. Un profil plus discret aurait été préférable sur un sujet aussi sensible », affirme-t-il. Pour le premier citoyen de Grugliasco, Roberto Montà, il s'agit par contre d'un faux problème. Franchement, je crois qu'il y a des questions plus importantes sur l'incinérateur qu'une comète. Et je trouve dangereuse l'instrumentalisation qui est faite de la question : elle ne sert qu'à effrayer les gens ».

Dans ces exemples, l'adverbe *francamente* sert à mettre en contraste des points de vue divergents : dans (138), le locuteur exprime son désaccord avec les commentateurs étrangers qui accusent les Italiens de dramatiser, soulignant au contraire la nécessité d'agir avec prudence. Dans (139), l'opinion de Monsieur Piazza est mise en contraste avec celle de Monsieur Montà concernant le positionnement d'une décoration en forme d'étoile sur un incinérateur. Bien que *penso che* puisse également être utilisé dans ce type de contextes (*francamente, penso che non possiamo più correre dietro a comparazioni di questa natura ; francamente, penso che sull'inceneritore ci siano questioni più importanti di una cometa*), il est notable que la combinaison privilégiée de *francamente* se construit de façon prototypique avec *credo che* et que *penso che* ne systématise pas ce procédé.

Les associations avec les CS de la classe *Evaluation*

Bien que pour la classe *Evaluation* *credo che* et *penso che* ne présentent pas de CS partagés de façon systématique dans les deux corpus, nous constatons qu'ils ont néanmoins la propriété commune d'introduire des jugements relevant de l'axiologique. *Credo che*, en particulier, présente plusieurs adjectifs et adverbes partagés dans PRES-IT et REP, à savoir *importante* (important), *giusto* (juste), *bene* (bien), *molto* (très, beaucoup), *davvero* (vraiment, très) :

- (140) Comunque siano entrati in Italia, comunque siano entrati i loro genitori in Italia, i loro diritti sono gli stessi di quelli nati in Italia e dei figli di italiani. **Credo che** sia molto importante crescere in questa convinzione, crescere con questa educazione. (PRES-IT)

Quelle que soit la manière dont ils sont entrés en Italie, quelle que soit la manière dont leurs parents sont entrés en Italie, leurs droits sont les mêmes que ceux des personnes nées en Italie et des enfants d'Italiens. Je crois qu'il est très important de grandir dans cette conviction, de grandir avec cette éducation.

- (141) Io credo di aver avuto qualche parte, da Presidente della Camera, per il varo della legge della riforma del 1993 che diede luogo alla elezione diretta dei sindaci e anche alla successiva scelta autonoma da parte dei sindaci dei componenti delle giunte, dei componenti degli organi di governo, e **credo che** sia stata una riforma giusta, una riforma buona. (PRES-IT)

Je crois avoir participé, en tant que président de la Chambre des députés, à l'adoption de la loi de réforme de 1993 qui a donné lieu à l'élection directe des maires et au choix autonome ultérieur par les maires des membres des conseils, des membres des organes de gouvernement, et je crois que c'était une réforme juste, une bonne réforme.

- (142) È vero che le avevano proposto di fare il vice sindaco e lei ha rifiutato? “**Credo che** Sveva Belviso ricopra l'incarico molto bene. Lei è una politica il mio è un ruolo tecnico”. (REP)

Est-il vrai qu'on vous a proposé d'être adjoint au maire et que vous avez refusé ? « Je crois que Sveva Belviso occupe très bien ce poste. C'est une femme politique, mon rôle est technique ».

- (143) **Credo che** questa sia davvero una questione da approfondire molto seriamente perché apre degli orizzonti diversi anche all'intervento dei poteri pubblici a tutela dei più deboli. (PRES-IT)

Je crois qu'il s'agit vraiment d'une question à examiner très sérieusement parce qu'elle ouvre des horizons différents aussi pour l'intervention des autorités publiques dans la protection des plus faibles.

Dans chacun de ces exemples, le locuteur met en scène des jugements de valeur. Ainsi, dans l'exemple (140), il souligne la nécessité d'une éducation à la tolérance en recourant à l'adjectif *importante*, renforcé par l'adverbe *molto*. Dans l'exemple (141), il qualifie la réforme de *giusta* et insiste sur son bien-fondé en employant l'adjectif *buona*. Dans l'exemple (142), il met en avant les compétences de Madame Belviso à travers l'expression *molto bene*. Enfin, dans l'exemple (143), il insiste sur l'importance d'approfondir une thématique donnée avec sérieux. Décrire une situation comme importante, juste, ou visant le bien, permet au locuteur

de s'inscrire dans un discours conforme à un système de valeurs socialement accepté et de se présenter ainsi comme une personne fiable, en adéquation avec ces normes.

Les associations avec les CS de la classe *Modalité*

Pour la classe *Modalité*, le verbe *potere* (pouvoir) est le premier marqueur qui apparaît en situation de cooccurrence statistiquement significative avec *credo che* et *penso che*. Cette attraction se retrouve dans les deux corpus :

- (144) Non so se siamo più forti dell'anno scorso. So solo che siamo forti. Voglio vincere tutto. **Penso che possiamo** vincere sia scudetto che Champions. (REP)

Je ne sais pas si nous sommes plus forts que l'année dernière. Je sais seulement que nous sommes forts. Je veux tout gagner. Je pense que nous pouvons gagner autant le *Scudetto* que la Ligue des champions.

- (145) Non è ancora altrettanto così nel penale – e infatti su questo tema il Procuratore Generale ha espresso delle parole preoccupate – ma **penso che** anche nel penale si possa avere un miglioramento in futuro. (PRES-IT)

Ce n'est pas encore le cas en droit pénal – le procureur général a d'ailleurs tenu des propos inquiétants à ce sujet – mais je pense qu'en droit pénal aussi, il peut y avoir une amélioration à l'avenir.

- (146) Ho visto, poi, che dal Libano è partito un appello per una cessazione del fuoco. **Credo che** l'Europa possa fare anche di più. Mi pare che il passo della Presidenza finlandese sia significativo e utile. (PRES-IT)

J'ai vu alors qu'un appel au cessez-le-feu venait du Liban. Je crois que l'Europe peut faire encore plus. Il me semble que la démarche de la Présidence finlandaise est significative et utile.

- (147) LUCAS SODDISFATTO – Una sensazione confermata dal suo allenatore: “**Credo che possa** nuotare molto più velocemente. Ha fatto comunque una performance corretta, non nuotava da parecchi mesi su questa distanza sulla vasca lunga, si può essere contenti di come sta andando”, ha detto Philippe Lucas. (REP)

LUCAS SATISFAIT – Une sensation confirmée par son entraîneur : « Je crois qu’il peut nager beaucoup plus vite. Il a quand même fait une bonne performance, ça faisait plusieurs mois qu’il n’avait pas nagé cette distance en bassin long, on peut être content de son parcours », a déclaré Philippe Lucas.

- (148) È irresponsabile anche la sola evocazione di ulteriori tagli. E dimostra quanto sia grave la deriva del governo Letta-Alfano. **Penso che** non si possa più colpire il welfare, la protezione sociale, i diritti dei cittadini. (REP)

La simple évocation de nouvelles réductions est irresponsable. Et cela montre la gravité de la dérive du gouvernement Letta-Alfano. Je pense qu’on ne peut plus toucher au bien-être, à la protection sociale, aux droits des citoyens.

- (149) Naturalmente, tutti i partiti, tutte le fasi storiche hanno luci ed ombre, ma **credo che** noi non possiamo dimenticare il contributo che quel grande partito, di ispirazione cristiana e di impegno democratico ha dato alla vita e allo sviluppo dell’Italia. (PRES-IT)

Bien sûr, tous les partis, toutes les phases historiques ont des lumières et des ombres, mais je crois que nous ne pouvons pas oublier la contribution que ce grand parti d’inspiration chrétienne et d’engagement démocratique a apportée à la vie et au développement de l’Italie.

- (150) “Credo che siano stati portati fin qua dalle correnti – spiega –. Qua attorno ci sono molte spiagge, anche libere, ma tutte sono attrezzate e controllate da bagnini. Nella giornata il mare era calmo, ma l’acqua era fredda. **Penso che** possano aver avuto una congestione, forse si sono gettati in mare dopo aver mangiato troppo. Prima dell’arrivo del 118 abbiamo fatto massaggio cardiaco e respirazione bocca a bocca”. (REP)

« Je crois qu’ils ont été rejetés ici par les courants – explique-t-il –. « Par ici, il y a beaucoup de plages, y compris des plages gratuites, mais toutes sont équipées et surveillées par des maîtres-nageurs. Ce jour-là, la mer était calme, mais l’eau était froide. Je pense qu’ils ont pu avoir une congestion, peut-être se sont-ils jetés à la mer après avoir trop mangé. Avant l’arrivée du 118, nous avons pratiqué le massage cardiaque et le bouche-à-bouche ».

Pour chaque corpus, dans au moins 90% des emplois où le verbe *potere* est cooccurrent de *credo che* et *penso che*, l’interprétation est radicale, avec des emplois de type capacité, possibilité matérielle ou permission. Lorsque *potere* exprime une capacité, « l’origine de la

possibilité réside dans les qualités inhérentes au sujet » (Barbet 2012 : 50) : c'est le cas des exemples (144), (146) et (147), où on met en avant les capacités d'une équipe de football de gagner plusieurs tournois (144), celles de l'Europe d'agir davantage (146) et celles d'un athlète de nager plus rapidement (147). Lorsque *potere* exprime une possibilité matérielle, la possibilité « demeure dans les conditions matérielles extérieures au sujet » (Barbet 2012 : 50) : c'est le cas de l'exemple (145). Enfin, il existe des emplois où le champ de la modalité du possible exprimée par *potere* « recouvre partiellement celui de la modalité du nécessaire » (Le Querler 1996 : 115), comme dans les emplois (148) et (149), où le locuteur exprime des sollicitations à ne pas mettre en péril les droits des citoyens et à ne pas oublier la contribution importante du parti Democrazia Italiana pour le développement du pays⁴⁹. Enfin, les interprétations épistémiques de *potere*, telles que celle présentée dans (150), restent très marginales dans les deux corpus. En somme, que ce soit une interprétation radicale ou épistémique du verbe *potere*, le fait d'introduire ce verbe modal par les séquences *credo che* ou *penso che* sert au locuteur à prendre de la distance d'un propos qui est décrit soit comme une possibilité (et qui a donc des chances de ne pas se réaliser), soit comme une obligation de ne pas faire (et qui peut donc être perçu comme menaçant pour l'interlocuteur).

Bien que *credo che* et *penso che* aient un comportement très semblable vis-à-vis du verbe *potere*, comme nous venons de le montrer, un motif linguistique propre à *credo che* émerge néanmoins dans le corpus des allocutions présidentielles. En effet, lorsque la séquence *credo che...potere* est mobilisée, on constate une forte tendance statistique à ce que le verbe complément du modal soit *dire*, comme en témoignent les exemples suivants :

- (151) Per alcuni aspetti **credo che si possa dire** che il tempo del forzato isolamento, del silenzio, talvolta della solitudine, abbia sollecitato tutti a porre in discussione il nostro rapporto con la realtà. (PRES-IT)

À certains égards, je crois qu'on peut dire que le temps de l'isolement forcé, du silence, parfois de la solitude, nous a tous poussés à nous interroger sur notre rapport à la réalité.

- (152) Abbiamo anche rapporti importanti in campo industriale, in campo scientifico e tecnologico: è veramente un modello di relazioni, quello

⁴⁹ Les emplois de *potere* de type « permission » sont systématiquement niés dans nos corpus. Pour Narrog (2012 : 188), une permission niée équivaut, d'un point de vue pragmatique, à une obligation de ne pas faire.

che si è stabilito tra l'Italia e l'Austria. E **credo che si possa dire** che abbiamo un comune sentire europeo. (PRES-IT)

Nous avons également d'importantes relations dans le domaine industriel, dans le domaine scientifique et technologique : c'est vraiment un modèle de relations qui a été établi entre l'Italie et l'Autriche. Et je crois qu'on peut dire que nous avons un sentiment européen commun.

(153) Non resteremo mai indifferenti alle stragi e alle sofferenze che si consumano nel Mare Nostro. E **credo che**, a testa alta, **possiamo dire** che l'Italia sta affrontando con dignità e grande umanità il proprio difficile compito. (PRES-IT)

Nous ne resterons jamais indifférents aux massacres et aux souffrances qui ont lieu dans Notre Mer. Et je crois que, la tête haute, nous pouvons dire que l'Italie affronte sa difficile tâche avec dignité et une grande humanité.

Dans PRES-IT, *dire* est le premier cooccurrent de la séquence *credo che...potere*⁵⁰, avec un indice de cooccurrence qui s'élève à 16,73. Comme le montrent les exemples, cette séquence s'actualise soit au moyen d'une forme impersonnelle (*credo che si possa dire che*), soit avec un pronom de première personne du pluriel (*credo che possiamo dire che*). La première séquence couvre le 59% des emplois et la deuxième le 24%. Dans le corpus journalistique, nous retrouvons également cette attraction, bien qu'avec un indice de cooccurrence moins élevé (4,15). L'analyse qualitative permet toutefois de montrer que dans REP cette séquence ne couvre que 20% des emplois où *credo...potere* et *dire* cooccurrent⁵¹, ce qui nous amène à conclure que ce motif linguistique n'est pertinent que pour le corpus des allocutions présidentielles. Ce résultat est en ligne avec la visée argumentative du type de textes qui composent ce corpus, ainsi qu'avec leur vocation à être préparés à l'avance pour être ensuite déclamés. Ce motif, nous le rappelons, particularise *credo che* par rapport à *penso che* : ce dernier ne s'insère pas, d'un point de vue statistique, dans ce type de construction spécifique.

⁵⁰ Les recherches ont été menées à partir du pivot [word="credo|Credo"] [itlemma="che"] []? []? []? []? []? []? []? []? [itlemma="potere" & verbe="verbe"] dans un empan droit de 5 items.

⁵¹ Le reste des emplois ne sont pas pertinents : soit il s'agit de cas où le verbe *croire* est nié, soit d'emplois où les verbes *croire* et *dire* ne forment pas de motif linguistique de la même manière que dans PRES-IT (comme dans l'exemple suivant : *Ho fiducia e credo che il nostro gruppo possa ancora dire tanto*).

Si nous nous concentrons à nouveau sur les marqueurs de la classe *Modalité*, nous remarquons qu'il y a deux autres marqueurs qui sont partagés par *credo che* et *penso che* dans les deux corpus : ce sont les verbes modaux *dovere* (devoir) et *bisognare* (falloir). Pour *credo che* uniquement, nous retrouvons également dans les deux corpus le substantif *bisogno* (besoin). Ces marqueurs servent à exprimer des jugements normatifs⁵², comme l'illustrent les exemples suivants :

- (154) Ne abbiamo parlato anche ieri con il Rettore, che mi ha accennato ai progetti che in questo senso si stanno delineando anche da parte della stessa Università. In quel caso **credo che** si debba, anche con l'impegno del governo nazionale, compiere ogni sforzo necessario per evitare che vada dispersa una straordinaria ricchezza ed una caratteristica veramente importante dell'economia veronese: mi riferisco a quella concentrazione di personale e di capitale umano altamente qualificato e impegnato nella ricerca di base. (PRES-IT)

Nous en avons également parlé hier avec le recteur, qui m'a fait part des projets que l'université elle-même est en train d'élaborer dans ce sens. Dans ce cas, je crois qu'on doit tout mettre en œuvre, y compris avec l'engagement du gouvernement national, pour éviter la dispersion d'une richesse extraordinaire et d'une caractéristique vraiment importante de l'économie véronaise : je parle de cette concentration de personnel hautement qualifié et de capital humain qui se consacre à la recherche fondamentale.

- (155) [...] gli elettori debbono effettivamente poter scegliere tra i candidati chi diano loro più fiducia. E **credo che** bisogna ridurre la frammentazione politica che in Italia ha raggiunto negli anni scorsi delle punte abnormi. (PRES-IT)

Les électeurs doivent en effet pouvoir choisir parmi les candidats en qui ils ont le plus confiance. Et je crois qu'il faut réduire la fragmentation politique qui a atteint des sommets anormaux en Italie ces dernières années.

⁵² Le verbe *dovere* est polysémique, tout comme son pendant français : un tri pour différencier les interprétations radicales des interprétations épistémiques a donc été effectué dans les deux corpus. Les interprétations épistémiques de *dovere* ne dépassent pas le 10% des occurrences lorsqu'il apparaît à côté de *credo che* ou *penso che*.

- (156) E un altro grosso problema è constatare che i soldi della Regione vengano utilizzati per produrre campagne senza idee. **Credo che** ci sia bisogno di ripensare le politiche delle Apt. (REP)

Un autre grand problème est le fait de constater que l'argent de la région est utilisé pour produire des campagnes sans idées. Je crois qu'il est nécessaire de repenser les politiques des Apt⁵³.

- (157) Si è fatto anche cenno al problema, che molto vi coinvolge, del rinnovo del contratto di lavoro: **penso che debba** essere normale, in un Paese civile, che i contratti di lavoro si rinnovino tempestivamente alla loro scadenza. (PRES-IT)

On a également évoqué le problème, qui vous concerne beaucoup, du renouvellement des contrats de travail : je pense qu'il doit être normal, dans un pays civilisé, que les contrats de travail soient renouvelés rapidement lorsqu'ils arrivent à échéance.

- (158) A Manchester abbiamo una forte attenzione nei confronti del settore giovanile – ha ricordato Ferguson – mentre in Italia ci si concentra maggiormente sulla prima squadra. Io **penso che bisogna** partire dalle fondamenta, ossia dai giovani, per creare un forte senso di appartenenza al club ed ottenere fiducia e fedeltà. (REP)

A Manchester, nous mettons l'accent sur le secteur des jeunes – rappelle Ferguson – alors qu'en Italie, nous nous concentrons davantage sur l'équipe première. Moi je pense qu'il faut partir de la base, c'est-à-dire des jeunes, pour créer un fort sentiment d'appartenance au club et gagner la confiance et la loyauté.

Dans tous ces exemples, les marqueurs *dovere*, *bisognare*, *bisogno* introduisent des jugements qui se rapprochent du normatif. En effet, ils présentent comme souhaitables (et donc à accomplir) des faits qui concernent des sujets d'actualité politique ou sociétale : dans les exemples (154) à (156), ces suggestions sont introduites par *credo che* et dans les exemples (157) à (158) par *penso che*. Dans (154), le locuteur suggère, à travers le verbe *dovere*, que l'Université de Vérone mette en œuvre des mesures qui puissent encourager les chercheur·euses à rester dans ses murs. Ses propos sont par ailleurs renforcés à travers d'autres marqueurs déontiques et évaluatifs : en qualifiant l'effort demandé de *necessario* (nécessaire),

⁵³ Aziende di Promozione Turistica (APT) = entreprises de promotion du tourisme.

il souligne l'inévitabilité de l'action à entreprendre et en décrivant le personnel comme une *straordinaria ricchezza* (richesse extraordinaire) et une *caratteristica veramente importante* (caractéristique vraiment importante), il valorise le capital humain et intellectuel comme un atout essentiel à préserver. Dans les exemples (155) et (156), les demandes de réduction de la fragmentation politique en Italie et de changement des politiques des agences de promotion touristique (APT) sont explicitées respectivement par le verbe *bisogna* et la locution *c'è bisogno di*. Dans (157), le locuteur exprime une évaluation personnelle sur une obligation sociale perçue comme évidente (*penso che debba essere normale*), en avançant ainsi une critique implicite du fait que les renouvellements des contrats de travail ne se déroulent pas de manière opportune. Enfin, dans (158), il suggère de se focaliser sur les footballeurs dès leur plus jeune âge afin de créer un esprit d'équipe fort.

Dans tous ces exemples, *credo che* et *penso che* fonctionnent comme des préfaces permettant d'introduire un contenu potentiellement dangereux pour l'interlocuteur, en raison de l'ordre qu'il contient (*si deve, bisogna, c'è bisogno di*). On observe que la fonction de *credo* en tant qu'écran de protection pour le locuteur, mise en relief par Cignetti (2009) dans des corpus oraux, se retrouve également dans des textes écrits. On peut toutefois se demander, dans l'exemple (158), si l'on n'est pas plutôt face à un usage attributif, comme le locuteur emploie le pronom *io* pour contraster son point de vue avec un point de vue précédent. Même avec un contexte élargi, il est souvent difficile d'établir une ligne très nette entre ces deux emplois de *penso che* et *credo che*.

5.3.2.2 *Sono convinto/a che et ritengo che*

Nous analysons dans cette section les verbes *sono convinto/a che* et *ritengo che* : les deux ont un profil combinatoire qui se rapproche de celui de *penso che* et *credo che*, car tous les deux ont tendance à attirer des CS des quatre classes sémantiques dans le deux corpus (la seule exception concerne la classe *Evaluation* pour *ritengo che* dans PRES-IT). Toutefois, des caractéristiques propres à ces verbes peuvent être mises en relief, notamment leur comportement face aux marqueurs de la classe *Modalité*.

Un premier point de convergence entre *sono convinto/a che* et *ritengo che* d'un côté et *penso che* et *credo che* de l'autre, est l'attraction statistique avec le pronom personnel *io* (classe *Prise en charge*). Ce résultat montre que les mêmes stratégies de focalisation sont mises en œuvre avec ces quatre verbes :

- (159) “Non mi interessa la celebrità per la celebrità – tiene a sottolineare – so bene che è come essere un acrobata sul filo, è una situazione pericolosa. **Io poi sono convinta che** dopo Millennium tutti mi odino; anzi, a ogni mio film che esce ho la sensazione dell’odio della gente...”. (REP)

« La célébrité pour la célébrité ne m’intéresse pas », souligne-t-elle, « je sais bien que c’est comme être un acrobate sur une corde raide, c’est une situation dangereuse. Et moi je suis convaincue qu’après Millennium, tout le monde me déteste ; en fait, à chaque film qui sort, je ressens la haine des gens... ».

- (160) Siamo due paesi fondatori, forse siamo, tra gli stessi paesi fondatori, quelli che hanno avuto nel corso dei decenni più coerenza e continuità nel sostenere la causa dell’unità europea; e ne siamo tuttora convinti sostenitori. **Io ritengo che** possiamo senz’altro sollecitare la massima intensificazione dei rapporti tra i nostri governi. (PRES-IT)

Nous sommes deux pays fondateurs, nous sommes peut-être, parmi les pays fondateurs eux-mêmes, ceux qui ont été les défenseurs les plus cohérents et les plus continus de la cause de l’unité européenne au cours des décennies ; et nous sommes toujours des défenseurs convaincus de cette cause. Je pense que nous pouvons sans aucun doute encourager la plus grande intensification des relations entre nos gouvernements.

Un autre point de convergence entre ces deux verbes, qui les rapproche également de *penso che* et *credo che*, est l’attraction constante à des connecteurs, avec une préférence systématique pour *perché* :

- (161) E non importa il ruolo che di volta in volta hanno le une o le altre, quali sono al governo e quali sono all’opposizione, **perché ritengo che** quando parliamo di competitività, e quindi in definitiva di futuro del sistema Paese, bisogna assolutamente trovare la strada per alcune fondamentali scelte condivise e per alcune politiche volte precisamente a elevare la competitività del sistema [...]. (PRES-IT)

Et peu importe le rôle que l’un ou l’autre joue de temps en temps, qui est au gouvernement et qui est dans l’opposition, parce que je considère que lorsque nous parlons de compétitivité, et donc en fin de compte de l’avenir du système national, il faut absolument trouver la voie de certains choix fondamentaux partagés et de certaines politiques visant précisément à accroître la compétitivité du système [...].

- (162) “Penso che l’arte e la musica nel mio caso, abbia una ruolo importante nel risvegliare le coscienze – ha detto – Ci metterò tutto il mio impegno e il mio cuore, perché sono convinta che la salute dei popoli sia uno dei primi antidoti contro la guerra [...]”. (REP)

« Je pense que l’art, et la musique dans mon cas, a un rôle important à jouer dans l’éveil des consciences – a-t-elle déclaré – J’y mettrai tout mon effort et mon cœur, parce que je suis convaincue que la santé des peuples est l’un des premiers antidotes contre la guerre [...] ».

Dans ces exemples, *perché* vient établir une relation argumentative entre une proposition *P* (*E non importa il ruolo che di volta in volta hanno le une o le altre, quali sono al governo e quali sono all’opposizione ; Ci metterò tutto il mio impegno e il mio cuore*) et la proposition *Q* qui contient le verbe d’indication de point de vue (*ritengo che quando parliamo di competitività... ; sono convinta che la salute dei popoli sia uno dei primi antidoti contro la guerra*). Ce mouvement argumentatif permet au locuteur de justifier les contenus dans *P* en se prédiquant la propriété d’être la source du point de vue explicité dans *Q*.

En ce qui concerne la classe *Modalité*, *sono convinto/a che* et *ritengo che* présentent des sélections très précises de marqueurs modaux, ce qui les distingue sur ce point de *credo che* et *penso che*. Nous observons en effet que *ritengo che* s’associe systématiquement avec le verbe *dovere* et *sono convinto/a che* avec le verbe *potere*, comme l’illustrent les exemples suivants :

- (163) Queste sono realtà essenziali per l’Italia. La SISSA è una di queste, e io **ritengo che** noi dobbiamo avere sempre molto fortemente presente educazione, ricerca, alta formazione, come garanzia per il nostro avvenire. (PRES-IT)

Ce sont des réalités essentielles pour l’Italie. Le SISSA est l’une d’entre elles, et moi je considère que nous devons toujours avoir à l’esprit l’éducation, la recherche, l’enseignement supérieur comme garantie de notre avenir.

- (164) “Penso che dovremo rivedere un po’ il sistema di reclutamento degli studenti all’estero – prosegue il nuovo direttore ipotizzando test che la Normale può organizzare anche fuori dall’Italia – oggi gli allievi stranieri sono il 10 per cento, io vorrei arrivare al 30 perché **ritengo che**

debba esserci una forte circolazione e un interscambio culturale nella formazione e nella ricerca”. (REP)

« Je pense que nous devons revoir un peu le système de recrutement des étudiants à l'étranger – poursuit le nouveau directeur, en émettant l'hypothèse de tests que la Normale peut aussi organiser en dehors de l'Italie – aujourd'hui les étudiants étrangers sont 10%, j'aimerais atteindre 30% parce que je considère qu'il doit y avoir une forte circulation et un échange culturel dans la formation et la recherche ».

- (165) Nel primo tempo, sul primo controllo trovava spesso il raddoppio di marcatura – spiega Prandelli – Abbiamo provato a farlo tagliare dal centro, e ci è riuscito: mi ha fatto piacere, i ragazzi credono a quel che stiamo facendo. Ma **sono convinto che** Antonio possa fare molto di più di quanto ha dato stasera. (REP)

En première mi-temps, sur le premier contrôle, il a souvent trouvé le double marquage – explique Prandelli – Nous avons essayé de lui faire couper depuis le centre et il a réussi : cela m'a fait plaisir, les garçons croient en ce que nous faisons. Mais je suis convaincu qu'Antonio peut faire beaucoup plus que ce qu'il a donné ce soir.

- (166) **Sono convinto che** possiamo limitare le conseguenze economiche e sociali della crisi mondiale per l'Italia, e creare anzi le premesse di un migliore futuro, se facciamo leva sui punti di forza e sulle più vive energie di cui disponiamo. (PRES-IT)

Je suis convaincu que nous pouvons limiter les conséquences économiques et sociales de la crise mondiale pour l'Italie, et même créer les prémisses pour un avenir meilleur, si nous tirons parti des forces et des énergies les plus vives dont nous disposons.

Le choix d'accompagner un jugement normatif par un verbe comme *ritengo che* (qui exprime un point de vue fondé sur une réflexion intellectuelle raisonnée⁵⁴) a pour effet de venir appuyer la demande de faire contenue dans *p* en l'étiquetant comme découlant d'un raisonnement réfléchi du locuteur. Les interprétations de *ritengo* dans (163) et dans (164) sont donc attributives et sont enclenchées par ailleurs par l'utilisation du pronom *io* (163) et du connecteur *perché* (164). Notons que le fait que *ritengo che* est particulièrement sensible à la modalité déontique converge avec les résultats de Strik Lievers (2012 : 163), qui relève son

⁵⁴ Cf. la définition de *considérer* dans le dictionnaire en ligne Larousse : « être d'avis après réflexion que » et celle de *ritenere* dans le dictionnaire en ligne Treccani : « Giudicare, stimare » (juger, estimer).

association privilégiée avec l'adjectif *necessario*. Le verbe *sono convinto/a che* a, de son côté, une préférence statistique pour le verbe *pouvoir* : dans les exemples (165) et (166), le but du locuteur est de se mettre en scène en se décrivant comme quelqu'un ayant confiance dans les capacités des autres.

5.3.2.3 *Mi pare che, sono sicuro/a che, dico che*

Les échelles de mise en scène énonciative du niveau macro (cf. section 5.1) ont permis de mettre en évidence le comportement très instable, d'un corpus à l'autre, de trois verbes, à savoir *mi pare che*, *sono sicuro/a che* et *dico che*. Rappelons en effet que *mi pare che* est spécifique à la macro-catégorie des CS dans PRES-IT mais pas dans REP ; inversement, *sono sicuro/a che* et *dico che* sont spécifiques à la macro-catégorie des CS dans REP mais pas dans PRES-IT. L'analyse au niveau micro permet de mieux comprendre ces résultats en faisant ressortir des régularités de comportement entre PRES-IT et REP pour *mi pare che* et de reconfirmer les fortes différences d'un corpus à l'autre pour *sono sicuro/a che* et *dico che*.

Mi pare che

En ce qui concerne *mi pare che*, des similarités de comportement entre PRES-IT et REP peuvent être mises en relief pour les classes *Prise en charge* et *Modalité* : en effet, ce verbe manifeste de manière très systématique une absence complète de marqueurs de ces deux classes. Pour la classe *Prise en charge*, le fait qu'on ne retrouve pas le pronom personnel *io* est justifié par l'agrammaticalité de la séquence *io mi pare che*⁵⁵. Toutefois, rien ne justifie, en théorie, l'absence d'attraction statistiquement significative avec des adverbes dans le contexte gauche de *mi pare che*, d'autant plus que nous observons quelques occurrences dans les deux corpus :

- (167) “Quello di gennaio per noi non sarà un mercato di riparazione: onestamente mi pare che non dobbiamo riparare nulla”, ha detto ieri Marotta, lasciandosi però scappare che “le problematiche maggiori sono a centrocampo”. (REP)

⁵⁵ La focalisation pourrait s'opérer au moyen du phénomène qu'on appelle *accusativo preposizionale* (cf. Treccani), qui donnerait la suite *a me (mi) pare che*. Néanmoins, les occurrences de *a me pare che* sont très rares dans nos corpus (2 dans PRES-IT et 8 dans REP) et aucune attestation de *a me mi pare che* n'a été relevée – ce qui n'est guère étonnant, puisque cet usage est uniquement répandu à l'oral informel.

« Le mois de janvier ne sera pas un marché de réparation pour nous : honnêtement, il me semble que nous n'avons rien à réparer », a déclaré hier Marotta, laissant toutefois entendre que « les plus gros problèmes se situent au milieu de terrain ».

- (168) Effettivamente, **mi pare che** si tratti di un programma imponente.
(PRES-IT)

En effet, il me semble qu'il s'agit d'un programme impressionnant.

Bien que les marqueurs de *Prise en charge* soient compatibles avec *mi pare che*, aucun cooccurrent spécifiques n'apparaît pour cette classe : ce résultat différencie *mi pare che* de tous les autres verbes du paradigme italien. Le fait que ce verbe marque un engagement très faible de la part du locuteur quant à la prise en charge du contenu introduit pourrait expliquer l'absence de marqueurs qui, au contraire, accentueraient sa voix dans l'énoncé.

La classe *Modalité* se caractérise également par l'absence de marqueurs, ce qui ne serait pas non plus prévisible sans une approche statistique, car *mi pare che* n'a pas de restrictions dans l'introduction des modalités du possible et du nécessaire, comme nous l'observons dans les exemples suivants :

- (169) Partendo dal presupposto che si tratti, a quanto mi si dice, di un bando non sufficientemente specifico, con tante questioni ancora da definire, pur nel rispetto dei rappresentanti delle istituzioni **mi pare che** la ricerca si possa ritenere almeno prematura. (REP)

En admettant qu'il s'agisse, comme on me l'a dit, d'un appel insuffisamment précis, avec de nombreuses questions encore à définir, tout en respectant les représentants des institutions, il me semble que la recherche peut au moins être considérée comme prématurée.

- (170) La necessità per Padalino sono scelte coerenti da parte del Parlamento: “Su questo **mi pare che** il legislatore debba prendere coscienza di quanto sta succedendo” ragionava Padalino. (REP)

Pour Padalino, il faut des choix cohérents de la part du Parlement : « Sur ce point, il me semble que le législateur doit prendre conscience de ce qui se passe », raisonnait Padalino.

- (171) Qui **mi pare che** non bisogna cedere ad allarmismi e vittimismo.
(PRES-IT)

Il me semble ici qu'il ne faut pas céder à l'alarmisme et à la victimisation.

Dans ces exemples, *mi pare che* accompagne des jugements de valeur (169) ou normatifs (170), (171), en marquant une atténuation de la force assertive de *p*. On pourrait dès lors s'attendre à ce que ce verbe joue un rôle similaire à celui de *penso che* et *credo che* dans la préparation de la suite du discours, en venant se combiner avec des modalités du possible et du nécessaire pour les affaiblir. Certes, cette stratégie discursive est tout à fait mobilisable par *mi pare che*, mais notre étude montre qu'elle n'est pas employée de façon systématique comme pour *penso che* et *credo che*. Cela illustre une fois de plus les avantages d'utiliser de manière complémentaire les méthodes de la linguistique théorique et les outils plus récents de la linguistique de corpus.

Sono sicuro/a che

Pour la forme *sono sicuro/a che*, le niveau macro indique que dans REP il est spécifiquement associé à la macro-catégorie des CS, et le niveau micro révèle la présence d'au moins un marqueur pour chaque classe sémantique. En revanche, dans PRES-IT, cette spécificité fait défaut : *sono sicuro/a che* n'est pas en association spécifique avec la macro-catégorie des CS et ne compte que deux marqueurs pour deux classes sémantiques, avec des indices relativement faibles. Cette rareté de marqueurs indique que, dans PRES-IT, *sono sicuro/a che* n'est pas utilisé de manière statistiquement significative dans des contextes où le locuteur met en scène de façon redondante sa propre voix. L'analyse qualitative des occurrences révèle que *sono sicuro/a che* apparaît principalement dans des contextes de prédiction, comme le montrent les exemples suivants :

- (172) Per finire, se mi consentite un accenno che mi è particolarmente caro, **sono sicuro che**, in questo spirito, dal Veneto verrà il giusto contributo alla celebrazione non retorica e non acritica del 150° Anniversario dell'Unità d'Italia. (PRES-IT)

Enfin, si vous me permettez une allusion qui m'est particulièrement chère, je suis sûr que, dans cet esprit, la Vénétie apportera la bonne contribution à la célébration non rhétorique et non acritique du 150e anniversaire de l'unification de l'Italie.

- (173) La Grecia ospiterà ad Atene, nel 2004, i Giochi Olimpici. L'Italia si sente vicina e solidale con voi per il successo di quest'impegno. **Sono sicuro che** il mondo sportivo italiano fornirà ogni possibile collaborazione al successo dei Giochi. (PRES-IT)

La Grèce accueillera, à Athènes en 2004, les Jeux Olympiques. L'Italie se sent proche et solidaire de vous pour la réussite de cette entreprise. Je suis sûr que le monde sportif italien apportera toute sa possible collaboration au succès des Jeux.

- (174) Vi ringrazio per quello che fate. E **sono sicuro che**, anche quest'anno, per tanti Italiani l'incontro con voi sarà occasione di crescita, di riflessione, di maggiore conoscenza e responsabilità. (PRES-IT)

Je vous remercie pour ce que vous faites. Et je suis sûr que, cette année encore, pour de nombreux Italiens, la rencontre avec vous sera l'occasion de grandir, de réfléchir, de mieux connaître et d'assumer des responsabilités.

Dans ces exemples, *sono sicuro/a che* est employé pour exprimer des états supposés se réaliser dans le futur, soutenus par des ancrages spatio-temporels tels que *150° Anniversario dell'Unità d'Italia* (173), *ad Atene, nel 2004* (173), ou *anche quest'anno* (174). Ces usages ne sont pas seulement fréquents, mais ils caractérisent statistiquement *sono sicuro/a che*. En effet, cette expression présente un indice de cooccurrence de 46,84 avec des verbes conjugués au futur (dans un empan droit de 10 items), tandis qu'aucune autre catégorie morphosyntaxique ne se révèle être spécifique à *sono sicuro/a che* dans ce corpus.

Dico che

Le dernier verbe qui affiche un comportement très divergent d'un corpus à l'autre est *dico che*. Nous rappelons en effet que cette expression est spécifique à la macro-catégorie des CS dans REP mais ne l'est pas dans PRES-IT. Les niveaux intermédiaire et micro montrent que dans REP ce verbe est sensible aux quatre classes sémantiques et présente pour chacune un ou plusieurs marqueurs, alors que dans PRES-IT nous ne relevons aucun marqueur pour aucune des classes. Les exemples suivants illustrent quelques emplois dans REP :

- (175) “[...] Io fumo il sigaro ma di fronte alla sanità sono disposto a pagare di più il mio vizio”. Il capo della Lega ha ammesso che si tratta di un “viziaccio” per cui brontolano sua moglie e i figli e che i medici gli hanno detto di evitare “ma io – ha concluso – **dico che** di qualcosa bisogna morire.” (REP)

« [...] Moi je fume le cigare, mais face à la santé, je suis prêt à payer plus pour mon vice ». Le chef de la Lega a admis qu'il s'agissait d'un « mauvais vice » dont sa femme et ses enfants se plaignaient et que les médecins lui avaient dit d'éviter, « mais moi – a-t-il conclu – je dis qu'il faut bien mourir de quelque chose ».

(176) Io dico che si può sognare pure il primo posto, però è difficile. (REP)

Moi je dis qu'on peut rêver de la première place, mais c'est difficile.

(177) “Sognare non costa nulla, noi dobbiamo puntare a essere tra i primi 5 e poi, con un pizzico di suspense, sarebbe bello non conoscere fino all'ultimo il finale – dichiara ammiccante De Laurentiis – personalmente comunque **dico che** è meglio la Champions che lo scudetto.” (REP)

« Rêver ne coûte rien, nous devons viser le top 5 et ensuite, avec un soupçon de suspense, ce serait bien de ne pas connaître le final jusqu'à la toute fin » – dit De Laurentiis en faisant un clin d'œil – personnellement je dis néanmoins que la Ligue des Champions vaut mieux que le *Scudetto*. »

(178) Mihajlovic non cerca giustificazioni: “Io dico che abbiamo imposto bene il nostro gioco e siamo riusciti a costruire occasioni importanti. È solo mancato un pizzico di cattiveria a Jovetic, Silva e Cerci per dare un corso diverso alla gara”. (REP)

Mihajlovic ne cherche pas d'excuses : « Moi je dis que nous avons bien imposé notre jeu et que nous avons réussi à créer des occasions importantes. Il nous a juste manqué un peu de méchanceté de la part de Jovetic, Silva et Cerci pour donner une autre tournure au match ».

Dans ces exemples, on retrouve des marqueurs qui sont en spécificité positive avec *dico che* dans REP : *ma* (*Construction argumentative*), *io* et *personalmente* (*Prise en charge*), *potere* et *bisognare* (*Modalité*) et *meglio* (*Evaluation*). On observe donc que le profil combinatoire de *dico che* dans le corpus de presse nationale se rapproche beaucoup de celui de *penso che* et *credo che*.

Ces résultats contrastent avec les résultats dans le corpus des allocutions présidentielles, où aucun CS n'a été relevé à côté de *dico che*. De plus, une analyse des exemples dans ce corpus montre que, dans 77% des cas, les emplois de *dico che* ne sont pas des emplois transparents :

(179) Al rappresentante dei detenuti, che ringrazio egualmente per le parole di apprezzamento rivoltemi, **dico che** considero importante la volontà di dialogo con le istituzioni da lui espressa, e la forte tensione di cui si è fatto portatore sul tema del reinserimento nel mondo del lavoro, per

non ricadere – come ha detto – nel circuito delinquenziale o scivolare nell'emarginazione. (PRES-IT)

Au représentant des détenus, que je remercie également pour les paroles d'appréciation qui m'ont été adressées, je dis que je considère comme importante la volonté de dialogue avec les institutions qu'il a exprimée, et la forte pression qu'il a exercée sur la question de la réinsertion dans le monde du travail, afin de ne pas retomber – comme il l'a dit – dans le circuit de la délinquance ou de ne pas glisser dans la marginalisation.

- (180) Più in generale le **dico che** ho ribadito e ribadisco quotidianamente che gli obiettivi fissati con il decreto di agosto, che prevedono in sostanza il raggiungimento del pareggio di bilancio entro il 2013 anziché il 2014 debbono essere assolutamente confermati anche alla fine del dibattito e delle votazioni in Parlamento. (PRES-IT)

Plus généralement, je peux vous dire que j'ai répété et que je répète quotidiennement que les objectifs fixés dans le décret du mois d'août, qui prévoient essentiellement un équilibre budgétaire pour 2013 au lieu de 2014, doivent absolument être confirmés, même à l'issue du débat et des votes au Parlement.

- (181) Quando **dico che** sono anche per noi una forza, cosa intendo dire? Che noi viviamo la realtà di un Paese che tende demograficamente a ridursi. È il problema delle cosiddette "culle vuote". (PRES-IT)

Quand je dis qu'ils sont aussi une force pour nous, qu'est-ce que je veux dire ? Que nous vivons la réalité d'un pays qui a tendance à se rétrécir démographiquement. C'est le problème des « berceaux vides ».

- (182) E quando **dico che** l'Italia ha sostenuto questo processo, intendo riferirmi ai governi assai diversi fra di loro che si sono succeduti alla guida del nostro Paese. (PRES-IT)

Et quand je dis que l'Italie a soutenu ce processus, je me réfère aux gouvernements très différents qui se sont succédé à la tête de notre pays.

Nous remarquons que dans ces exemples, *dico che* ne remplit pas les critères de verbe transparent : il n'est pas utilisé pour introduire un contenu et si on le supprime, une partie de l'information communiquée est perdue. Dans les exemples (179) et (180), *dico che* est utilisé pour s'adresser aux interlocuteurs par la formule *A (x) dico che...* ; dans (181) et (182), il permet au locuteur de revenir sur son propre discours pour reformuler et mieux expliquer ses

propos. Ces usages relèvent de routines discursives propres au genre des allocutions, où le président s'adresse oralement à son public et met en évidence volontairement son acte de parole. On voit donc bien la différence avec les emplois (175) à (178), où *dico che* se rapproche davantage de la communication d'un point de vue, notamment dans le cadre d'interviews sportives.

5.3.3 Synthèse du niveau micro

L'analyse au niveau micro a d'abord mis en évidence la similitude de comportement entre *penso che* et *credo che* dans les deux corpus étudiés. Ces deux expressions apparaissent principalement dans des contextes où le point de vue du locuteur est mis à l'avant-plan discursif, souvent en opposition avec d'autres point de vue, et sont fréquemment associées à des adjectifs axiologiques. En outre, *credo che* et *penso che* accompagnent fréquemment des jugements normatifs introduits par des verbes modaux tels que *dovere* et *bisognare*. Nous avons également relevé un motif linguistique spécifique à *credo che* (*credo che si possa dire che*), caractéristique du corpus des allocutions présidentielles. Quant à *ritengo che* et *sono convinto/a che*, leur profil combinatoire se rapproche de celui de *credo che* et *penso che*, même s'ils effectuent des sélections précises par rapport aux marqueurs de la classe *Modalité* (*ritengo che...dovere* ; *sono convinto/a che...potere*). Enfin, l'étude des trois verbes présentant des comportements distincts dans les deux corpus a permis de mettre en lumière des usages spécifiques de deux d'entre eux dans PRES-IT : les emplois prédictifs de *sono sicuro/a che* et les emplois de *dico che* en tant que commentaire mettant en évidence le caractère performatif renvoyant à l'acte de parole.

6. Mise en perspective des résultats français et italiens

Ce chapitre propose une analyse comparative de la combinatoire des verbes d'indication de point de vue en français et en italien, dans le but de mettre en évidence les similitudes et les divergences entre ces deux langues. Les tendances combinatoires propres à chacun des deux paradigmes verbaux, identifiées dans les chapitres 4 et 5, servent désormais de base à cette comparaison interlinguistique. Cette dernière sera menée en procédant par niveau d'analyse (macro, intermédiaire et micro).

6.1 Les niveaux macro

Au niveau macro, une hiérarchie similaire se dessine pour les deux langues. En effet, on observe une dichotomie entre certains verbes qui se situent tout en haut de l'échelle de mise en scène énonciative (il s'agit des verbes qui ont une attraction statistique forte avec la macro-catégorie des CS, comme en témoignent les valeurs de *log-likelihood*), et d'autres verbes qui se trouvent au pôle opposé et qui ne sont pas spécifiques à la macro-catégorie.

En français, cette opposition est particulièrement régulière : dans les deux corpus analysés, les expressions *je crois que* et *je pense que* occupent le pôle positif des deux échelles, tandis qu'en bas figurent *je suis sûr·e que* et *je suis convaincu·e que*. En italien, nous retrouvons les mêmes verbes en haut des deux échelles (*credo che* et *penso che*), mais les tendances sont moins nettes pour ce qui concerne les verbes non spécifiques. En effet, dans le corpus des allocutions présidentielles nous observons les verbes *sono sicuro/a che* et *dico che* et dans le corpus de presse nationale *mi pare che*. Une différence très nette que nous pouvons mettre en relief entre français et italien, est le comportement de *je suis convaincu·e que/sono convinto/a che* : la forme française n'est pas spécifique, dans les deux corpus, à la macro-catégorie des CS, alors que la forme italienne apparaît au troisième rang juste après *credo che* et *penso che*.

Remarquons enfin que les indices de *log-likelihood* des verbes italiens sont, en moyenne, presque cinq fois plus élevés que ceux des verbes français. Ce phénomène est particulièrement notable dans le corpus de presse REP, où ces indices reflètent la forte présence énonciative du locuteur dans les environnements dans lesquels les verbes italiens apparaissent.

6.2 Les niveaux intermédiaires

L'analyse du niveau intermédiaire permet de mieux comprendre la composition interne des macro-catégories associées à chaque verbe. Nous remarquons que plus les verbes d'indication de point de vue sont proches du pôle positif de l'échelle, plus les différents types de CS qui entrent en jeu sont nombreux : pour les verbes tout en haut de l'échelle, les quatre classes sémantiques sont représentées. Cette observation est valable pour les deux langues : *je crois que, je pense que, credo che, penso che* présentent pour chaque corpus une attraction aux quatre classes sémantiques. Les classes qui affichent les indices les plus hauts sont la classe *Modalité* (en français et en italien) ainsi que la classe *Prise en charge* (en italien). *Sono convinto/a che* est également sensible aux quatre classes, bien qu'avec des indices moins hauts (rappelons que son correspondant morphologique français *je suis convaincu-e que* n'est soit spécifique à aucune classe dans PRES-FR, soit à une uniquement dans LM).

Parmi les verbes qui ne présentent pas de spécificité à l'égard de la macro-catégorie de CS, plusieurs configurations sont observées :

1. Certains verbes ne sont attirés par aucune classe sémantique : c'est le cas de *je suis convaincu-e que* dans PRES-FR, de *il me semble que* dans LM, et de *dico che* dans PRES-IT) ;
2. Certains verbes sont attirés par une seule classe sémantique : *je suis sûr-e que* est attiré par la classe *Evaluation* dans PRES-FR et par la classe *Construction argumentative* dans LM ; *je suis convaincu-e que* est attiré par la classe *Construction argumentative* dans LM ; *mi pare che* est attiré par la classe *Construction argumentative* dans REP ;
3. Le verbe *sono sicuro/a che* présente une attraction pour deux classes sémantiques dans PRES-IT, à savoir *Prise en charge* et *Evaluation*.

Ces tendances sont assez compréhensibles. Lorsque l'indice de la macro-catégorie est de 0, aucune classe n'est représentée, car aucun CS a été repéré (cas de figure 1) ; mais lorsque la valeur de *log-likelihood* se rapproche du seuil de spécificité (10,83), cela indique que des CS ont été observés, ce qui peut influencer la spécificité d'une ou plusieurs classes sémantiques (cas de figure 3 : la valeur d'attraction entre *sono sicuro/a che* et sa macro-catégorie des CS dans PRES-IT est de 10,42).

6.3 Les niveaux micro

Cette section s'appuie sur les résultats des analyses des deux niveaux micro, qui ont permis d'élaborer les profils combinatoires des verbes en français (section 4.3) et en italien (section 5.3), fondés sur les cooccurrents spécifiques qui les accompagnent. L'analyse qui suit adopte une approche contrastive visant à mettre en perspective le fonctionnement des formes morphologiquement ou sémantiquement proches dans les deux langues, en identifiant les points saillants qui les rapprochent ou les distinguent.

Nous commencerons par examiner les verbes présentant des profils combinatoires très similaires en français et en italien, indépendamment du corpus dans lequel ils apparaissent : *je crois que/credo che, je pense que/penso che, je considère que/ritengo che, il me semble que/mi pare che*. Ensuite, l'analyse se concentrera sur les verbes dont les comportements semblent diverger dans les deux langues : *je suis convaincu·e que/sono convinto/a che, je suis sûr·e que/sono sicuro/a che, et je dis que/dico che*.

Le Tableau 6 fusionne les résultats des tableaux récapitulatifs pour les verbes français et italiens (Tableau 4 et Tableau 5). Chaque verbe français est comparé à son équivalent morphologique italien ou au verbe italien le plus proche sémantiquement. Nous excluons de ce tableau *je trouve que*, car il n'a pas de correspondant dans le paradigme des verbes italiens. L'interprétation suit la même logique que pour les Tableau 4 et Tableau 5. Une case verte indique la présence d'au moins un marqueur commun dans les deux langues (que nous soulignons), ou l'absence partagée de marqueurs (le symbole « \emptyset » souligné) ; une case rouge signifie qu'aucun marqueur n'est partagé (ou que l'absence de marqueurs n'est pas partagée).

<i>Construction argumentative</i>	<i>Prise en charge</i>		<i>Modalité</i>	<i>Evaluation</i>
<u>mais</u>	<u>moi</u>	Je crois que	<u>falloir</u> <u>pouvoir</u> <u>devoir (v.)</u>	-
<u>ma</u> però	<u>io</u> francamente	Credo che	<u>potere</u> <u>dovere (v.)</u> <u>bisognare</u> bisogno	molto davvero bene importante giusto
<u>mais</u>	<u>moi</u> personnellement	Je pense que	<u>falloir</u> <u>pouvoir</u> <u>devoir (v.)</u>	<u>bon</u> utile assez
<u>ma</u> perché	<u>io</u>	Penso che	<u>potere</u> <u>dovere (v.)</u> <u>bisognare</u>	<u>buono</u>
<u>mais</u> <u>parce que</u>	-	Je considère que	<u>devoir (v.)</u>	-
<u>perché</u>	io	Ritengo che	<u>dovere (v.)</u>	-
∅	<u>∅</u>	Il me semble que	-	∅
-	<u>∅</u>	Mi pare che	∅	-
-	∅	Je suis convaincu·e que	∅	∅
<u>ma</u> perché	io	Sono convinto/a che	potere	grande
-	∅	Je suis sûr·e que	∅	-
-	io	Sono sicuro/a che	-	-
-	moi	Je dis que	falloir	-
-	-	Dico che	-	-

Tableau 6 : tableau comparatif des CS communs aux verbes d'indication de point de vue dans les deux langues.

6.3.1 Je crois que/credo che et je pense que/penso che

Ces quatre formes sont analysées de manière conjointe car leur profil combinatoire est très proche. Tout d'abord, elles présentent, pour les quatre classes, une richesse de marqueurs qui les distingue des autres verbes du paradigme ; de plus, elles partagent entre elles beaucoup de ces marqueurs et cela indépendamment du corpus dans lequel elles sont employées. Dans les pages qui suivent, nous mettons en relief les principaux points communs entre ces formes, mais aussi deux points de divergence qui ont été relevés.

- *Je crois que/credo che* et *je pense que/penso che* sont utilisés de façon statistiquement significative pour exprimer des évaluations personnelles, formulées au moyen d'adjectifs axiologiques qui peuvent à leur tour être modifiés par des adverbes de degré. Très systématique est en particulier la combinaison entre *je pense que/penso che* et l'adjectif *bon/buono*.
- *Je crois que/credo che* et *je pense que/penso che* sont mobilisée de façon préférentielle dans des contextes où le locuteur porte son point de vue à l'avant-plan discursif, au moyen du connecteur *mais/ma*. De plus, la posture énonciative du locuteur se voit renforcée par l'attraction statistiquement significative de ces formes avec le pronom *moi/io* ou des adverbes dans le contexte gauche. Les sélections préférentielles sont toutefois différentes entre français et italien : en français, c'est l'adverbe *personnellement* qui se combine de façon statistiquement significative avec *je pense que*, alors qu'en italien c'est l'adverbe *francamente* qui vient accompagner *credo che*.
- Ces formes sont utilisées de façon statistiquement significative pour introduire des jugements normatifs exprimés par des verbes modaux tels que *falloir/bisognare* et *devoir/dovere*. Ainsi, *je crois que/credo che* et *je pense que/penso che* ne se limitent pas à l'expression d'évaluations personnelles, mais participent à des stratégies discursives qui visent à préserver la face de l'interlocuteur, particulièrement lorsqu'il s'agit de formuler des demandes ou des recommandations susceptibles d'être perçues comme menaçantes.
- Enfin, l'analyse du profil combinatoire de ces expressions met en lumière un type particulier de motif discursif construit autour du verbe *croire/credere* : il s'agit des expressions *je crois qu'on peut dire que/credo che si possa dire che* qui interviennent de façon statistiquement significative dans les deux corpus des allocutions présidentielles. On peut dès lors avancer l'hypothèse qu'il ne s'agit pas d'un simple

motif récurrent, mais d'une véritable « routine », au sens défini par Sitri et Tutin (2016). En effet, cette notion est évoquée pour

qualifier des séquences lexico-syntaxiques « prêtes à écrire », étroitement associées à des pratiques discursives spécifiques, liées à des communautés bien délimitées, la communauté scientifique ou la communauté des éducateurs par exemple. Ces façons de dire sont étroitement liées aux genres institués (Maingueneau, 2004) et jouent un rôle important dans l'identité sociale des scripteurs. C'est en partie en s'appropriant ces routines que les scripteurs se définissent comme appartenant à une communauté de discours donnée. (Sitri et Tutin 2016 : 9-10).

Nous considérons qu'il est pertinent de parler de routine dans ce cas précis, car cette séquence s'est systématisée de manière identique dans les deux langues, formant une *séquence lexico-syntaxique* « prête à écrire ». De plus, elle est étroitement associée aux discours présidentiels, qui sont des textes à visée argumentative, destinés à être oralisés, et dans lesquels le locuteur s'inscrit et se met en scène explicitement.

6.3.2 Je considère que/ritengo che

Les expressions *je considère que* et *ritengo che* sont ici abordées conjointement en raison de leurs profils combinatoires remarquablement similaires. En effet, dans les deux corpus pris en compte, pour les deux langues, elles montrent une attraction systématique avec les mêmes cooccurrents, à savoir *parce que/perché* et *devoir/dovere*. Ces marqueurs peuvent apparaître ensemble, comme nous l'observons dans les exemples suivants :

(183) Vous savez, pour qu'il y ait de l'emploi, il faut que nous créions les conditions, et je fais en sorte qu'elles soient là, y compris à travers le Pacte de responsabilité, c'est-à-dire la baisse des charges et la baisse des cotisations – parce que je considère que les entreprises doivent avoir les moyens d'investir et d'embaucher – mais deuxièmement, il faut qu'il y ait de la croissance, de la croissance. (PRES-FR)

(184) Ho ricevuto un sms nel pomeriggio da amici del movimento No Tav, sono andato lì per vedere e filmare con il cellulare perché ritengo che la popolazione debba essere informata sul passaggio di convogli speciali che trasportano scorie nucleari. (REP)

J'ai reçu un SCS dans l'après-midi de la part d'amis du mouvement No Tav, je suis allé là-bas pour voir et pour filmer avec mon téléphone portable car je considère que la population doit être informée du passage des convois spéciaux transportant des déchets nucléaires.

Les propriétés sémantiques intrinsèques à ces deux verbes (*je considère que* et *ritengo che* expriment un point de vue fondé sur une réflexion intellectuelle raisonnée) influencent leur profil combinatoire. En effet, l'attraction statistique avec *parce que/perché* peut être interprétée comme une conséquence des traits sémantiques qui les caractérisent : lorsque le locuteur se trouve à devoir justifier un contenu donné (*je fais en sorte qu'elles soient là/sono andato lì per vedere et filmare*) au moyen d'un verbe d'indication de point de vue, il privilégiera un verbe qui a une valeur attributive (comme *je considère que/ritengo che*), plutôt qu'un verbe qui a uniquement une valeur modale (comme *il me semble que/mi pare che*). L'association entre *je considère que/ritengo che* et le verbe modal *devoir/dovere* s'inscrit dans le même type de stratégie argumentative : le fait de préfacer *devoir/dovere* par *je considère que/ritengo che* contraint l'interprétation radicale du verbe modal (qui exprime une obligation) et permet au locuteur d'augmenter la force du but illocutoire de la complétive.

6.3.3 Il me semble que/mi pare che

Il me semble que et *mi pare che* sont, de manière systématique dans les deux corpus et dans les deux langues, les expressions les moins sensibles aux marqueurs des quatre classes sémantiques (cf. Tableau 4 et Tableau 5). Une hypothèse pour expliquer ce comportement peut être le fait que les propositions construites au moyen de *sembler/sembrare* et *paraître/parere* présentent systématiquement un jugement épistémique qui constitue une conclusion provisoire d'un raisonnement (Kratschmer 2006). Le locuteur, en utilisant ces expressions, tendrait à prendre moins directement en charge les contenus qu'il énonce, réduisant ainsi sa responsabilité quant à la validité des informations transmises. L'analyse qualitative des occurrences de *il me semble que/mi pare che* dans nos corpus révèle que ces expressions peuvent s'inscrire dans des contextes où interviennent d'autres marqueurs contribuant à une prise de distance vis-à-vis de *p*. Considérons les exemples suivants :

(185) Je ne la connais pas assez bien⁵⁶, mais **il me semble qu'**il y a une schizophrénie entre le rêve américain et la nouvelle réalité américaine qui s'est cristallisée au tout début du XXe siècle [...]. (LM)

(186) Non vorrei sbagliare i nomi, **mi pare che** siano Damian Czarnecki e Arianna Giulianelli. (PRES-IT)

Je ne voudrais pas me tromper sur les noms, je pense qu'il s'agit de Damian Czarnecki et Arianna Giulianelli.

Dans ces exemples, les propositions soulignées renforcent la prise de distance par rapport à *p* : dans (185), *je ne la connais pas assez bien* laisse ouverte la possibilité de remettre en question le contenu présenté par le locuteur (*il y a une schizophrénie entre le rêve américain et la nouvelle réalité américaine*) ; dans (186), la proposition *non vorrei sbagliare i nomi* modalise également le discours en signalant que le locuteur se montre prudent quant à l'affirmation qu'il émet. Cependant, ces mécanismes discursifs ne peuvent être systématiquement identifiés dans notre approche actuelle, car ils reposent sur des constructions complexes que l'étiquetage que nous utilisons ne permet pas de capturer. Pour approfondir ces observations et valider nos hypothèses, il serait nécessaire d'élaborer un étiquetage systématique de ces séquences discursives afin de déterminer si leur usage conjoint à *il me semble que/mi pare che* est statistiquement pertinent.

6.3.4 Je suis convaincu·e que/sono convinto/a che et je suis sûr·e que/sono sicuro/a che

L'analyse en trois niveaux a permis, pour le français, de mettre en relief un emploi prédictif de *je suis convaincu·e que* et *je suis sûr·e que* : ces expressions servent davantage à formuler des suppositions qu'à exprimer un jugement évaluatif. Cette caractéristique est illustrée par les exemples suivants :

(187) On ne peut pas dire que tous ces artisans sont artistes, mais il y a des artistes parmi eux, nuance Rashmi Poddar, rédactrice en chef de la revue d'art Marg. « **Je suis sûr·e que** bientôt, ces derniers entreront dans les musées d'art contemporain, comme les bronzes de l'ethnie Bastar qui sont déjà vendus très chers. » (LM)

⁵⁶ Le locuteur se réfère à la société américaine.

- (188) Mais enfin **je suis convaincu que** nous aurons au printemps, ou au début de l'été, la troisième phase de la Conférence, à moins d'événements extraordinaires... (PRES-FR)

Dans les exemples de nos corpus, les emplois prédictifs de *je suis convaincu-e que/je suis sûr-e que* s'accompagnent fréquemment d'adverbes temporels, tels que *que bientôt* (187), *au printemps, au début de l'été* (188). En italien, seul *sono sicuro/a che* adopte un comportement similaire, mais uniquement dans le corpus des allocutions présidentielles :

- (189) È ovvio che non spetta a me suggerire la risposta alla grossa domanda indirizzata a tutti noi: "Il Regno Unito e l'Italia: condividiamo lo stesso futuro?" **Sono sicuro che** le risposte verranno dalle sedute plenarie e dai workshop che si succederanno fino a domenica. (PRES-IT)

Il est évident qu'il ne m'appartient pas de suggérer la réponse à la grande question qui est adressée à nous tous : « Le Royaume-Uni et l'Italie : partageons-nous le même avenir ? » Je suis sûr que les réponses viendront des sessions plénières et des ateliers qui se poursuivront jusqu'à dimanche.

- (190) **Sono sicuro che** i prossimi due anni dimostreranno in maniera incontrovertibile come due Paesi, fondatori dell'Unione Europea, sapranno lavorare insieme per contribuire a sviluppare relazioni sempre migliori nella comunità internazionale. (PRES-IT)

Je suis sûr que les deux prochaines années démontreront de manière incontestable comment deux pays, fondateurs de l'Union Européenne, sauront travailler ensemble pour contribuer à développer des relations toujours meilleures au sein de la communauté internationale.

Comme son correspondant français, *sono sicuro/a che* introduit des prédictions concernant des événements futurs et cette lecture prédictive est accompagnée par des adverbiaux temporels comme *fino a domenica* (189) et *i prossimi due anni* (190). En revanche, *sono convinto/a che* n'affiche pas ce comportement exclusif. Bien que des emplois prédictifs soient présents, ce verbe se distingue par une attraction très variée avec divers marqueurs des quatre classes sémantiques, ce qui rapproche davantage son profil combinatoire à celui de *penso che* et *credo che*. Ces observations mériteraient d'être approfondies avec d'autres corpus pour confirmer la persistance de ces tendances en français et comprendre les raisons de cette différence concernant *sono convinto/a che*.

6.3.5 Je dis que/dico che

L'analyse des corpus français révèle une propriété particulière associée à *je dis que*, à savoir son association statistiquement significative avec *falloir*, comme celle qui intervient dans l'exemple suivant :

- (191) Qui dit partage laisse entendre qu'il y aurait une quantité donnée de travail et qu'il faudrait la partager plus équitablement entre ceux qui en ont et ceux qui n'en ont pas. Moi j'ai un raisonnement tout à fait différent. **Je dis qu'il faut** d'abord augmenter la quantité de travail disponible et que c'est à partir de là seulement que l'on peut raisonner.
(PRES-FR)

Dans cet extrait, le locuteur met en contraste différents points de vue. Il commence par exposer une position qu'il attribue à d'autres – celle d'un partage équitable de la quantité de travail disponible – tout en s'en distanciant par l'emploi du conditionnel (*il y aurait*). Par la suite, il oppose ce point de vue au sien à l'aide de deux procédés : d'une part, par l'énoncé *moi, j'ai un raisonnement tout à fait différent*, qui marque explicitement le contraste avec le point de vue précédent ; d'autre part, par l'introduction de son propre jugement normatif (*il faut d'abord...*) accompagné par *je dis que*.

En italien, aucune propriété stable partagée n'émerge pour *dico che* dans les deux corpus étudiés. Dans le corpus REP, *dico che* présente un profil combinatoire proche de celui de *credo che* et *penso che*, avec la présence récurrente de marqueurs spécifiques pour chacune des classes sémantiques. Dans le corpus PRES-IT, *dico che* ne présente aucun CS et l'analyse des occurrences a permis de mettre en relief des emplois propres à ce corpus, à savoir des emplois où *dico che* renvoie à l'acte de parole du locuteur, qui l'exploite soit pour s'adresser à son public, soit pour revenir sur ses propos et les reformuler. Ce type d'emploi semble fortement influencé par les textes composant le corpus des allocutions présidentielles, qui ont pour vocation d'être déclamées devant les citoyens.

6.4 Synthèse du chapitre comparatif

Dans ce chapitre, nous avons exploré les similitudes et différences entre les profils combinatoires des verbes d'indication du point de vue en français et en italien. L'analyse comparative a permis de relever que certains correspondants morphologiques, comme *je crois*

que/credo che et *je pense que/penso che*, présentent des profils combinatoires presque superposables, partageant une forte sensibilité aux quatre classes sémantiques et par conséquent des usages discursifs similaires. D'autres verbes, proches sémantiquement, affichent également des comportements superposables (*je considère que/ritengo che* et *il me semble que/mi pare che*). Enfin, nous avons pu mettre en relief le fait qu'une proximité morphologique entre deux verbes n'implique pas forcément un profil combinatoire semblable : *je suis convaincu-e que/sono convinto/a che* et *je dis que/dico che* révèlent des divergences d'emploi en fonction des langues et des corpus pris en compte.

7. Conclusion

Cette partie conclusive s'articule en trois sections. La section 7.1 revient sur les questions de recherche initiales et apporte des réponses globales, qui seront ensuite approfondies dans la section 7.2. Celle-ci met en lumière les principales contributions de la recherche, en proposant également des domaines d'application. Enfin, la section 7.3 présente les limites et explore les pistes de recherche futures.

7.1 Bref retour sur les objectifs

Dans cette étude, nous avons analysé un paradigme de verbes d'indication de point de vue, c'est-à-dire des formes verbales utilisées avec un morphème de première personne de l'indicatif présent, introductrices de propositions complétives et employées pour présenter un contenu comme relevant de la voix du locuteur. Ce type de verbes a fait, dans la littérature à ce sujet, l'objet d'une analyse approfondie dans leur usage en incise ; en revanche, leur emploi non-parenthétique a été moins étudié, sauf dans le cadre de la tradition logico-philosophique française. Des travaux comme ceux de Ducrot (1975), Carel (2011b), Tuchais (2012), Gosselin (2015) ont exploré les distinctions sémantiques entre verbes proches sémantiquement et ont mis en évidence des différences subtiles entre eux. L'objectif de cette thèse était d'évaluer si une approche quantitative pouvait compléter les descriptions sémantiques existantes, en tenant compte des contextes privilégiés dans lesquels ces formes sont utilisées. Commençons par apporter des réponses aux questions de recherche introduites au chapitre 1 :

- **Quels sont les contextes d'apparition privilégiés des verbes d'indication de point de vue ?**

Bien que tous les verbes d'indication de point de vue mettent explicitement en scène le locuteur dans son discours, en présentant le contenu qu'ils introduisent sur un ton engagé (Carel 2011a), leurs contextes d'apparition ne sont pas complètement superposables. En effet, certains verbes (comme *je pense que*, *je crois que*, *penso che*, *credo che*) ont tendance à apparaître dans des contextes où l'on retrouve beaucoup de marqueurs explicitant la voix du locuteur, alors que d'autres (comme *je suis sûr-e que*, *je suis convaincu-e que*, *dico che*, *mi pare che*) sont statistiquement moins sensibles à ces marqueurs.

- **Est-ce que ces contextes d'apparition privilégiés mettent en lumière des similarités et des différences sémantiques entre ces formes ?**

Oui, les contextes d'apparitions privilégiés permettent de mettre en lumière des similarités de comportement entre verbes sémantiquement proches, par exemple entre *je crois que* et *je pense que* en français, ainsi qu'entre *credo che* et *penso che* en italien. Mais ils permettent également de relever certaines différences, comme celle qui concerne le motif reliant *je pense que* à l'adverbe *personnellement*, qui ne se réalise pas avec *je crois que*. De plus, des motifs spécifiques ont été relevés pour certaines formes : *je dis que* s'associe de façon statistiquement significative au verbe *falloir* et *je considère que* au verbe *devoir*. Pour l'italien, nous observons également un motif reliant *ritengo che* au verbe *dovere*.

- **Est-ce que les propriétés observées pour chaque forme sont stables selon le type de corpus mobilisé ?**

Bien que le taux de représentation statistique des verbes d'indication de point de vue varie selon le corpus (les analyses des sur et sous-représentations montrent des tendances opposées entre les corpus des discours présidentiels et les corpus de presse nationale), leurs propriétés sémantiques sont, en règle générale, très stables entre les différents types de textes mobilisés. En effet, comme nous l'avons vu dans les Tableau 4 et Tableau 5, la majorité des verbes des deux paradigmes partagent un profil combinatoire très proche, présentant plusieurs marqueurs communs aux différents corpus, surtout en français. Les verbes d'indication de point de vue italiens *sono sicuro/a che* et *dico che* sont les formes les plus instables d'un corpus à l'autre.

- **Est-ce que les propriétés observées pour les verbes français se retrouvent pour les verbes morphologiquement ou sémantiquement équivalents en italien ?**

Oui, dans la plupart des cas, les verbes morphologiquement ou sémantiquement proches ont un profil combinatoire semblable en français en italien. Toutefois, des différences d'emplois entre *je suis convaincu·e que* et *sono convinto/a che*, ainsi qu'entre *dico che* et *je dis que* ont été relevées.

Dans la section suivante, nous apporterons des éléments de réponse supplémentaires à ces questions de recherche, tout en mettant en lumière les apports les plus marquants de ce travail.

7.2 Apport de la thèse

Les premières contributions apportées par cette thèse sont de type méthodologique :

- La démarche *corpus-driven* adoptée pour extraire le paradigme des verbes d'indication de point de vue a permis d'identifier un groupe de formes verbales caractérisées par un usage prototypique à la première personne du singulier, au présent de l'indicatif et introduisant une proposition complétive. Toutes les formes retenues ont un emploi transparent, c'est-à-dire qu'elles présentent le contenu de la complétive comme relevant de la voix du locuteur. Un résultat particulièrement intéressant est que l'application de la même méthodologie à des corpus français et italiens a permis de mettre en évidence deux paradigmes de verbes très similaires d'une langue à l'autre. Les analyses de sur et sous-représentation présentées dans la section 3.3.4 ont démontré que ces verbes forment une famille cohérente, avec des tendances statistiques nettes dans les deux langues : dans les corpus composés de discours présidentiels, ces verbes sont tous suremployés, alors que dans les corpus composés d'articles de presse nationale, ils sont en grand partie sous-employés.
- Le paradigme de verbes d'indication de point de vue n'est pas biaisé par des catégorisations sémantiques faites à priori, car il a été constitué sur la base de l'usage effectif de ces verbes dans les corpus sélectionnés. Ce choix méthodologique a favorisé la création d'un paradigme de formes qui ne sont pas traditionnellement considérées comme formant un tout. Par exemple, le verbe *dire* est généralement décrit comme un verbe de parole et n'est jamais associé aux mêmes paradigmes qui comprennent des verbes comme *penser*, *croire*, *trouver*. Cependant, nous estimons que des constructions comme *je pense que p*, *je crois que p* ou *je trouve que p* partagent des caractéristiques communes avec *je dis que p*, notamment le pouvoir de présenter un contenu *p* en l'étiquetant comme relevant de la voix du locuteur. Ces réflexions sont évidemment tout aussi valables pour les formes italiennes.

Les apports théoriques que nous présentons ci-dessous sont le fruit du dialogue entre les résultats quantitatifs et le retour au texte qualitatif. Plusieurs tendances stables entre les verbes d'indication de point de vue français et italiens ont pu être relevées :

- Les verbes *je pense que* et *je crois que*, ainsi que *penso che* et *credo che*, se distinguent comme les plus fréquents des deux paradigmes. De plus, ces quatre verbes ont un profil combinatoire très proche qui montre qu'ils ont tendance à apparaître dans des

environnements où l'on retrouve de nombreux marqueurs mettant en scène la voix du locuteur. On remarque en particulier une tendance nette de ces formes à être introduites par *mais/ma*, *moi/io* et à accompagner des évaluations à dimension axiologique. En français, on observe une dichotomie assez frappante entre les verbes qui ont tendance à apparaître dans des contextes où le locuteur fait entendre sa voix via les marqueurs des quatre classes (*je pense/crois que*) et les verbes qui ne sont aucunement sensibles à ces marqueurs (*je suis convaincu-e que* ou *je suis sûr-e que*). Cette différence pourrait s'expliquer par la nécessité pour *je pense que* et *je crois que* de mobiliser des marqueurs de mise en scène énonciative afin d'activer leur interprétation attributive (pensons, par exemple, aux cooccurrents *moi* et *personnellement*). Ceci ne serait pas nécessaire pour des formes telles que *je suis convaincu-e que* et *je suis sûr-e que*, car elles n'ont pas besoin de s'associer à ce type de marqueur pour véhiculer une interprétation attributive. Cette tendance de *je pense/crois que* à activer leur sens attributif via la mise en scène du pronom personnel *moi* et/ou de l'adverbe *personnellement* serait donc liée à des mouvements argumentatifs. Cette hypothèse pourrait être vérifiée en testant la force d'attraction du connecteur argumentatif *mais* avec les séquences présentant une dislocation à gauche (*moi/personnellement + je + verbe + que*) et en la comparant avec la force d'attraction de *mais* avec les mêmes séquences sans dislocation (*je + verbe + que*).

- Certaines préférences sémantiques précises ont été relevées pour certains verbes : en français, *je considère que* introduit de façon statistiquement significative des jugements normatifs exprimés par le verbe modal *devoir*, tandis que *je dis que* accompagne de façon privilégiée le verbe modal *falloir*. Lorsque ces deux verbes sont associés à des modalités déontiques, la force du but illocutoire de la proposition complétive est accrue. Cette stratégie se retrouve également en italien, notamment dans la combinaison de *ritengo che* avec le verbe *dovere*.
- Certaines formes peuvent avoir un usage spécifique selon le corpus dans lequel elles sont mobilisées. C'est notamment le cas de *je crois que* et *credo che* pour lesquels nous avons relevé la même routine discursive dans les deux langues : *je crois qu'on peut dire que p/credo si possa dire che p*. Par le biais de cette séquence, le locuteur module son engagement vis-à-vis de *p*. On observe ici la différence entre cet emploi, où le locuteur se dissimule derrière un dire indéterminé, et le motif *personnellement je pense que p*, par lequel il se met en scène en décrivant son raisonnement.

Les comportements pratiquement superposables entre formes morphologiquement proches que nous venons de discuter ne sont pas aléatoires. En effet, le fait que les mêmes tendances aient été observées, pour chaque langue, dans deux corpus différents, suggère que les propriétés mises en relief font intrinsèquement partie du sémantisme de ces verbes. Il serait toutefois nécessaire d'approfondir et confirmer ces observations en prenant en compte d'autres types de textes.

Passons enfin aux différences relevées entre le français et l'italien :

- Par rapport au français, l'italien codifie une plus forte présence de la voix du locuteur lorsque ces verbes sont mobilisés. En effet, dans les échelles du niveau macro, les indices de *log-likelihood* des verbes italiens sont, en moyenne, presque cinq fois plus élevés que ceux des verbes français ; de plus, au niveau micro, les listes de cooccurents spécifiques sont très riches, notamment dans le corpus de presse nationale. L'exemple de *sono convinto/a che* est à ce propos emblématique. On observe que ce verbe présente un comportement très différent de son correspondant morphologique *je suis convaincu-e que*, car son profil combinatoire le rapproche beaucoup plus des verbes *penso/credo che* et il n'accompagne pas de façon statistiquement significative un emploi prédictif (observé pour son pendant français). Une hypothèse pour cette différence pourrait être le fait que *sono convinto/a che* aurait davantage un sens modal qu'attributif, en ce sens qu'il serait utilisé principalement pour introduire un point de vue, sans pour autant attribuer au locuteur la propriété d'être *convinto/a che p*.
- L'autre différence remarquable concerne *dico che*. On observe qu'il ne partage absolument pas son profil combinatoire avec celui de son correspondant morphologique *je dis que* : le motif très constant *je dis que...falloir* ne se retrouve pas de manière systématique dans les deux corpus italiens. En italien, selon le corpus, les emplois de *dico che* varient : dans le corpus de presse, sont davantage présents des emplois modaux (*dico che* est utilisé, dans le cadre d'interviews sportives, pour communiquer un point de vue tout en l'atténuant), tandis que dans les discours présidentiels, il s'agit principalement d'un emploi d'adresse, très vraisemblablement influencé par le fait que le but des textes composant ce corpus soit de s'adresser oralement à un public.

7.2.1 Champs d'applications

Les apports mis en évidence dans la section précédente peuvent trouver plusieurs champs d'application. Tout d'abord, dans le domaine des études en sémantique, car les analyses effectuées permettent d'apporter des données quantitatives qui peuvent aider à définir les propriétés des verbes pris en compte. En particulier, l'utilisation d'outils statistiques joue un rôle clé en offrant la possibilité d'identifier des tendances générales et des régularités d'usage qui ne peuvent pas être détectées par une approche se basant sur des exemples construits.

Ce travail peut également contribuer aux avancées récentes en traduction automatique. À titre d'illustration, examinons un extrait tiré de notre corpus REP, traduit en français à l'aide de l'outil DeepL :

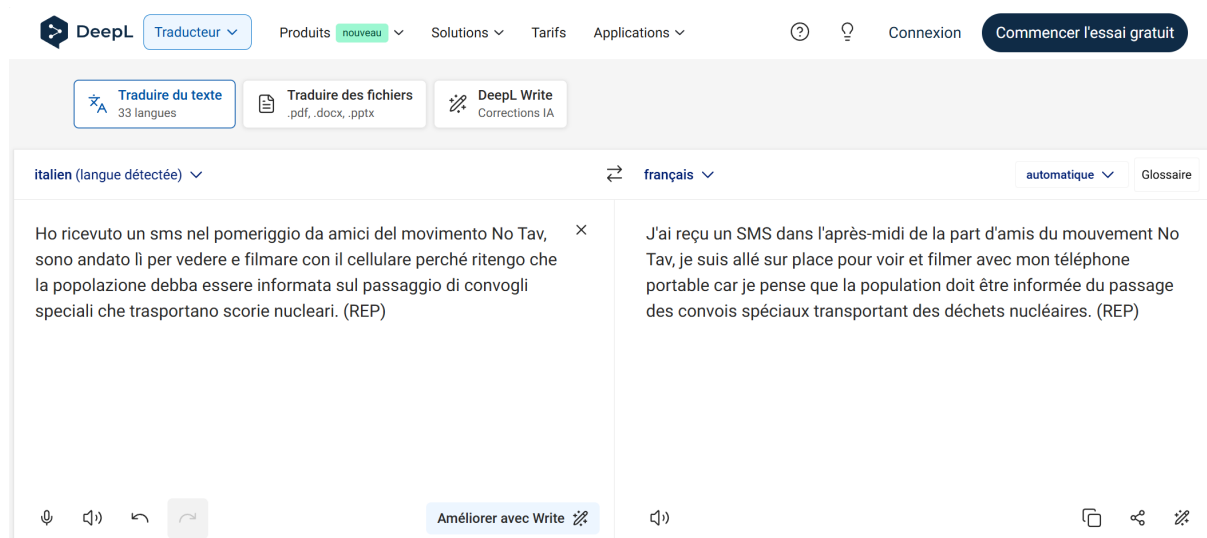


Figure 22 : exemple de traduction sur *DeepL*.

Dans cet exemple, le verbe *ritengo* est traduit par *je pense*, probablement en raison de critères de fréquence, puisque l'expression *je pense que* est très courante en français. Cependant, le choix de recourir à un verbe qui concède une interprétation modale, comme *penser*, pour accompagner le verbe *devoir*, a pour effet d'affaiblir la force du but illocutoire de *p*, alors qu'en italien le choix d'utiliser *ritenere* semble plutôt venir renforcer l'obligation présente dans *p*. Ainsi, pour gagner en précision, il serait conseillé de traduire par *car je considère que la population doit*.

Enfin, cette étude présente un intérêt particulier dans le domaine de l'enseignement-apprentissage des langues. Nos recherches peuvent notamment contribuer à améliorer les ressources proposées dans les manuels de FLE pour l'expression de l'opinion

personnelle. Delahaie et al. (2022) ont passé en revue les structures proposées dans les *Référentiels pour le français* en mettant en évidence les expressions recommandées pour chaque niveau :

Référentiel	Structures proposées	Exemples donnés
Référentiel A1 pour le français (2007)	(Moi,) <i>Je pense que p</i> <i>Je crois que p</i>	<i>Je pense que Julie a raison.</i>
Référentiel A2 pour le français (2008)	Les structures du niveau A1 + <i>Pour moi, p</i> (Moi) <i>je trouve que p</i> <i>À mon avis</i>	<i>Pour moi, il va mieux.</i> <i>Je crois que Julie est partie.</i>
Référentiel B1 pour le français (2011)	Niveau A1-A2 + <i>Il me semble que p</i> <i>D'après moi, p</i> <i>Personnellement, p</i>	<i>À mon avis, les manifestants ont raison.</i> <i>Il me semble que Jean va mieux.</i> <i>D'après moi, ça va mal se passer entre eux.</i> <i>Personnellement, ça ne me plaît pas trop.</i>
Référentiel B2 pour le français (2004)	Niveau A1-A2-B1+ <i>De mon point de vue, p</i> <i>En ce qui me concerne, p</i> <i>Pour / selon moi, p</i> (Moi,) <i>je considère que p</i> (Moi,) <i>personnellement, p</i> <i>Mon sentiment, c'est que p</i> <i>Si tu veux savoir ce que je pense, p</i>	<i>Selon nous, elle ne va pas rester longtemps célibataire.</i> <i>Mon sentiment, c'est qu'il aurait mieux fait de se taire.</i> <i>Si vous voulez savoir ce que je pense, je suis tout à fait d'accord.</i>

Figure 23 : Recensement des structures servant à introduire l'opinion personnelle dans les *Référentiels pour le français* (Delahaie et al., 2022)

Observons tout d'abord que ces listes n'incluent pas la totalité des verbes que nous avons inclus dans notre paradigme : il manque *je suis sûr-e que*, *je suis convaincu-e que* et *je dis que*. De plus, certains choix de classement par niveau peuvent être questionnés : pourquoi les expressions *pour/d'après/selon moi, p* sont introduites à des niveaux plus hauts que *je pense/crois que p* ? Et faut-il vraiment enseigner une expression telle que *Mon sentiment, c'est que p* qui semble assez artificielle ? Si oui, à quel(s) niveau(x) et dans quel(s) contexte(s) ? La prise en compte de l'usage réel de ces marqueurs peut aider à améliorer les ressources à disposition. Par exemple, une simple étude de leur fréquence permettrait de créer des listes des expressions les plus courantes et de problématiser la question de celles qui ne sont pas, ou très peu, attestées. D'autres choix de présentation de ces structures dans les *Référentiels* ne sont pas très convaincants. Par exemple, notre étude a montré comment, dans les corpus pris en compte, le pronom *moi* n'a pas de préférence entre *je pense/crois/trouve que*, contrairement à ce qui est suggéré dans le *Référentiel*, où l'on ne recommande pas son emploi avec *je crois que p*. Un autre exemple concerne l'adverbe *personnellement*. Dans le *Référentiel*, son utilisation est préconisée dans des structures telles que *Personnellement, p* et *Moi,*

personnellement, p, mais jamais comme introduction aux verbes d'opinion. La combinaison (*moi*) *personnellement je pense que p*, attestée dans nos corpus, est même rejetée par certain·es locuteur·rices qui la considèrent comme un « pléonasme » ou un « tic de langage insupportable »⁵⁷. Enfin, l'analyse comparative proposée dans cette étude est particulièrement pertinente pour des apprenant·es italophones étudiant le français. Elle présente en effet plusieurs avantages. Premièrement, elle peut contribuer à rassurer les apprenant·es en mettant en évidence des similarités importantes entre les deux langues, ce qui facilite l'établissement de correspondances entre la langue cible et leur langue maternelle. Deuxièmement, elle souligne les différences en démontrant que des formes morphologiquement superposables peuvent ne pas partager les mêmes traits sémantiques. Enfin, elle met en avant une dimension souvent négligée dans les manuels de FLE, à savoir le fait que ces verbes contribuent à la construction de stratégies d'atténuation et de gestion des faces, et que le choix d'un verbe plutôt qu'un autre peut produire des effets rhétoriques distincts.

7.3 Limites et prolongements

Cette étude présente certaines limites qui ouvrent des perspectives de recherche.

Tout d'abord, le paradigme étudié demeure restreint : le nombre de verbes d'indication de point de vue analysés pourrait être élargi, car de nombreux autres verbes mériteraient une attention similaire. Avec l'accès à des corpus de plus grande envergure, il serait envisageable d'explorer des marqueurs morphologiques italiens non pris en compte par notre méthodologie actuelle, tels que *considero che, trovo che, ou mi sembra che*. De même, il serait intéressant de comprendre le fonctionnement d'autres marqueurs d'introduction du point de vue du locuteur, comme les marqueurs « médiatifs » ou « évidentiels » (Coltier et Dendale 2004 : 41) *pour moi, selon moi et à mon avis*, en intégrant également leurs correspondants italiens (*per me, secondo me, a mio avviso*). Une autre piste d'investigation pourrait être l'étude de l'impact

⁵⁷ Voir notamment les discussions dans la section « Langue française » du journal *Le Figaro* en ligne : <https://www.lefigaro.fr/langue-francaise/expressions-francaises/2018/04/09/37003-20180409ARTFIG00003-les-expressions-a-bannir-au-bureau-moi-personnellement-je-pense.php>

<https://www.lefigaro.fr/langue-francaise/expressions-francaises/2019/01/27/37003-20190127ARTFIG00010-dix-pleonasmes-a-eradiquer-de-vos-conversations.php>

du pronom *on* dans l'introduction du point de vue : quels seraient les profils combinatoires de formes comme *on croit que*, *on pense que*, *on dit que* ?

Par ailleurs, intégrer d'autres corpus pourrait enrichir les résultats et mieux aider à déterminer si les tendances des verbes d'introduction du point de vue sont inhérentes à leur sémantisme ou influencées par le type de texte. La prise en compte d'un corpus de textes issus de blogs ou d'un recueil de commentaires publiés sur les réseaux sociaux pourrait fournir des données pertinentes pour approfondir l'analyse. Il serait également intéressant de s'intéresser aux textes générés automatiquement par les chabots tels que ChatGPT.

De plus, notre analyse a révélé que les marqueurs de prise en charge incluent souvent des adverbes de phrase qui apparaissent dans le contexte gauche de l'expression, comme *personnellement/personnellement*. Cependant, notre méthodologie actuelle n'a pas permis d'identifier les adverbes de prédicat qui pourraient également intervenir à proximité de ces verbes. Par exemple, les formes complexes telles que *je crois vraiment que* ou *je dis simplement que* mériteraient une investigation plus poussée. Dans un premier temps, il conviendrait de déterminer quels types d'adverbes cooccurrent avec quels prédicats (cette question a été partiellement abordée dans Sanvido 2024). Ensuite, il serait pertinent de vérifier si le profil combinatoire des formes complexes diffère de celui des formes simples correspondantes.

Enfin, notre étude s'est concentrée sur un empan phrastique, c'est-à-dire que nous avons limité notre analyse aux unités de sens contenues dans un seul énoncé composé d'une proposition *Q* et une proposition complétive *p*. Cette restriction a permis de mener une exploration approfondie à ce niveau, mais elle a également réduit la portée de nos observations. En élargissant la recherche à des unités textuelles plus étendues, comme le paragraphe ou le discours entier, il serait possible d'examiner d'autres marqueurs linguistiques. Ces marqueurs, situés en dehors de l'énoncé contenant le verbe d'indication de point de vue, pourraient fournir des informations complémentaires et enrichir l'analyse. Par exemple, pour les connecteurs, notre étude s'est concentrée sur ceux apparaissant immédiatement avant le verbe d'indication de point de vue (contexte gauche étroit), ce qui a permis d'identifier les relations directes entre le connecteur et le verbe. Cependant, une analyse portant sur les connecteurs qui suivent la proposition complétive (contexte droit élargi) serait également pertinente. Cela permettrait de mieux comprendre comment le discours s'organise après l'apparition de ce type de verbes et d'explorer les effets qu'ils produisent sur la structuration des énoncés qui suivent.

8. Bibliographie

8.1 Références principales

- Abeillé, A. & Koenig, J.-P. (2021). Les classes sémantiques des verbes. In Abeillé, A., Godard, D., Delaveau, A., Gautier, A. (eds.). *La Grande Grammaire du Français*, vol. 1. Arles : Actes Sud, 180-208.
- Aijmer, K. (1997). I think - an English modal particle. In Swan, T. & Westvik, O. J. (eds.). *Modality in Germanic languages. Historical and comparative perspectives*. Berlin/New York : Mouton de Gruyter, 1-47.
- Akissi Boutin, B. (2007). De, que, pour, subordonnants, et variation en français. *Linx*, 57, 57-68.
- Andersen, H. L. (1997). *Propositions parenthétiques et subordination en français parlé*. Thèse de Doctorat, Université de Copenhague.
- Andersen, H. L. (2007). Marqueurs discursifs propositionnels. *Langue française*, 2, 13-28.
- Angot, J. (2021). *Epistemic and subjective expressions in French : the case of je pense, je crois and je trouve*. Thèse de Doctorat, University of Manchester.
- Anscombe, J.-C. (2002). Mais/pourtant dans la contre-argumentation directe : raisonnement, généricité, et lexicque. *Linx. Revue des linguistes de l'université Paris X Nanterre*, 46, 115-131.
- Anscombe, J.-C. (2010). Autour d'une définition linguistique des notions de voix collective et de ON-locuteur. In Colas-Blaise, M., Kara, M., Perrin, L., Petitjean, A. (eds.). *La question polyphonique ou dialogique en sciences du langage*. Metz : Université Paul Verlaine, 39-64.
- Apothéloz, D. (2003). La rection dite « faible » : grammaticalisation ou différentiel de grammaticité ? *Verbum*, 25, 241-262.
- Avanzi, M. (2009). La prosodie des verbes parenthétiques en français parlé. *Linx. Revue des linguistes de l'université Paris X Nanterre*, 61, 131-144.
- Baroni, M., Bernardini, S., Comastri, F., Piccioni, L., Volpi, A., Aston, G., Mazzoleni, M. (2004). Introducing the La Repubblica Corpus. A Large, Annotated, TEI (XML)-compliant Corpus of Newspaper Italian. In *Proceedings of LREC 2004*.

- Barbet, C. (2012). Devoir et pouvoir, des marqueurs modaux ou évidentiels ? *Langue française*, 173(1), 49-63.
- Battaglia, E. & Lo Baido, M. C. (sous presse). *Evidential and epistemic functions between syntax and discourse : verbs of thought in spoken Italian*. In Pardal Padin, A. (ed.). *Verbs of thought and speech. Pragmaticalization paths across languages*. John Benjamin Publishing Company : Amsterdam/Philadelphia.
- Beeching, K. (2007). La co-variation des marqueurs discursifs bon, c'est-à-dire, enfin, hein, quand même, quoi et si vous voulez : une question d'identité ? *Langue française*, 154, 78-93.
- Begioni, L. (2013). Mode et temps dans les subordonnées complétives, étude comparative français-italien, le cas des verbes d'opinion tels que croire/credere, espérer/sperare, penser/pensare. *Studia UBB Philologia*, 58(4), 187-196.
- Benveniste, E. (1966). *Problèmes de linguistique générale, I*. Paris : Gallimard.
- Benzécri, J.-P. & coll. (1981). *Pratique de l'analyse des données : linguistique et lexicologie*. Paris : Dunod.
- Benzecri, J.-P. (1966). Linguistique et mathématique. *Revue Philosophique de La France et de l'Étranger*, 156, 309-374. URL : <http://www.jstor.org/stable/41090270>
- Benzécri, J.-P. (1982). *Histoire et préhistoire de l'analyse des données*. Paris : Dunod.
- Bernard, T. (2015). *Verbes d'attitude propositionnelle et analyse discursive*. Thèse de doctorat, Université Paris Diderot-Paris 7.
- Biber, D. (2015). Corpus-based and corpus-driven analyses of language variation and use. In Heine, B. & Narrog, H. (eds). *The Oxford handbook of linguistic analysis* [2nd edition]. Oxford University Press, 193-224.
- Biber, D., Johansson, S., Leech, G., Conrad, S., Finegan, E. (1999). *The Longman Grammar of Spoken and Written English*. New York : Longman.
- Blanche-Benveniste, C. & Willems, D. (2007). Un nouveau regard sur les verbes « faibles ». *Bulletin de la Société de Linguistique de Paris*, 102(1), 217-254.
- Blanche-Benveniste, C. & Willems, D. (2016). Les verbes faibles, *Encyclopédie Grammaticale du Français* [En ligne]. URL : http://encyclogram.fr/notx/010/010_Notice.php

- Blanche-Benveniste, C. (1989). Constructions verbales « en incise » et rection faible des verbes. *Recherches sur le français parlé*, 9, 53-73.
- Blumenthal, P. (2002). Profil combinatoire des noms. Synonymie distinctive et analyse contrastive. *Zeitschrift für französische Sprache und Literatur*, 112, 115-138.
- Blumenthal, P. (2008). Histoires de mots : affinités (s)électives. *Congrès mondial de linguistique française*, Paris, 31-46. URL : <https://doi.org/10.1051/cmlf08338>
- Blumenthal, P. (2012). Méthodes statistiques en lexicologie contrastive. In Begioni, L. & Bracquenier, C. (eds). *Sémantique et lexicologie des langues d'Europe. Théories, méthodes, applications*. Presses Universitaires de Rennes, 113-128.
- Borillo, A. (1976). Les adverbes et la modalisation de l'assertion. *Langue française*, 30, 74-89.
- Borillo, A. (1982). Deux aspects de la modalisation assertive : croire et savoir. *Langages*, 67, 33-53.
- Borillo, A. (2004). Les « adverbes d'opinion forte » selon moi, à mes yeux, à mon avis, ... : point de vue subjectif et effet d'atténuation. *Langue française*, 142, 31-40.
- Bronzi, A. M. (1977). Indicativo e congiuntivo nelle complete italiane. *Studi di Grammatica italiana*, 6, 425-450.
- Brown, P. & Levinson, S. C. (1990). *Politeness : Some universals in language usage* [2nd edition 1978/Reprint]. Cambridge university press.
- Brugger, G. & D'Angelo, M. (1995). Movement at LF triggered by mood and tense. *Folia Linguistica*, 29(3-4), 195-221.
- Busso, L. & Lenci, A. (2016). Italian VerbNet : A Construction-based Approach to Italian Verb Classification. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*, European Language Resources Association, Portorož, Slovenia, 2633-2642.
- Caffi, C. (1999). On mitigation. *Journal of Pragmatics*, 31, 881-909.
- Caffi, C. (2001). *La mitigazione. Un approccio pragmatico alla comunicazione nei contesti terapeutici*. Münster : LIT.
- Carel, M. & Ducrot, O. (2009). Mise au point sur la polyphonie. *Langue française*, 164(4), 33-43.
- Carel, M. (2023). *Parler*. Campinas : Pontes Editores.

- Carel, M. (2011a). La polyphonie linguistique. *Transposition* [En ligne]. URL : <http://journals.openedition.org/transposition/365>
- Carel, M. (2011b). *L'entrelacement argumentatif. Lexique, discours et blocs sémantiques*. Paris : Honoré Champion.
- Ceccaldi-Hamet, A. & Lacaze, G. (2020). Le discours rapporté et l'expression de la subjectivité. *e-Rea* [En ligne], 17.2. URL : <https://doi.org/10.4000/erea.10018>
- Cellier, P., Quiniou, S., Charnois, T., Legallois, D. (2012a). Fouille de données pour la stylistique : l'exemple des motifs émergents. In Dister, A., Longrée, D., Purnelle, G. (eds). *Actes des Journées Internationales d'analyse statistique des données textuelles (JADT12)*. Liège, Belgique, 821-833. URL : <http://lexicometrica.univ-paris3.fr/jadt/>
- Cellier, P., Quiniou, S., Charnois, T., Legallois, D. (2012b). What About Sequential Data Mining Techniques to Identify Linguistic Patterns for Stylistics? *Lecture Notes in Computer Science*, vol. 7181, 166-177.
- Cignetti, L. (2009). Sémantique di *credo* nello scritto e nel parlato. *Analisi corpus-driven*. In Ferrari, A. (ed). *Sintassi storica e sincronica dell'italiano: subordinazione, coordinazione, giustapposizione: atti del X Congresso della Società internazionale di linguistica e filologia italiana*, vol. 2. Firenze: Franco Cesati Editore, 939-954.
- Coltier, D. & Dendale, P. (2004). La modalisation du discours de soi : éléments de description sémantique des expressions pour moi, selon moi et à mon avis. *Langue française*, 2, 41-57.
- Cornulier, B. de (1978). L'incise, la classe des verbes parenthétiques et le signe mimique. *Cahiers de Linguistique de l'Université de Québec*, 8, 53-95.
- Cortelazzo, M. A. (1995). La lingua italiana di fine millennio. *Belfagor*, 50(1), 103-112.
- Dal Negro, S. (2015). Contatto linguistico e organizzazione del discorso : il ruolo dei verbi. In Consani, C. (ed). *Contatto interlinguistico fra presente e passato*. Milano : LED, 83-100.
- De Cesare, A.-M., Garassino, D., Agar Marco, R., Albom, A., Cimmino, D. (2016). *Sintassi marcata dell'italiano dell'uso medio in prospettiva contrastiva con il francese, lo spagnolo, il tedesco e l'inglese. Uno studio basato sulla scrittura dei quotidiani online*. Peter Lang.

- Delahaie, J., Rouanne, L., Canut, E. (2022). Les verbes recteurs faibles dans l'expression de l'opinion en français et en espagnol. *Travaux de linguistique*, 84-85, 131-157.
- Della Valle, V. & Patota, G. (2009). *Viva il congiuntivo ! Come e quando usarlo senza sbagliare*. Milano : Sperling & Kupfer.
- Dessaintes, M. (1960). *La construction par insertion incidente (étude grammaticale et stylistique)*. Paris : D'Artrey.
- Detges, U. & Waltereit, R. (2014). *Moi je ne sais pas vs. Je ne sais pas moi : French Disjoint Pronouns in the Left vs. Right Periphery*. In Beeching, K. & Detges, U. (eds.). (2014). *Discourse functions at the left and right periphery : Crosslinguistic investigations of language use and language change*. Leiden/Boston : Brill.
- Diwersy, S., Legallois, D. (2014). L'apport de la méthode des motifs aux analyses phraséologiques en discours. *Approches théoriques et empiriques en phraséologie* [En ligne], Nancy, France. URL : <https://hal.science/hal-03546271v1>
- Doro-Mégy, F. (2008). *Etude croisée de think, believe, croire et penser*. Paris : Éditions Ophrys.
- Dubois, J. & Dubois-Charlier, F. (1997). *Les verbes français*. Paris : Larousse-Bordas.
- Ducrot, O., Bourcier, D., Bruxelles, S., Diller, A. M., Fouquier, E., Gouazé, J., Maury, L., Nguyen, T. B., Nunes, G., Ragunet de Saint-Alban, L., Rémis, A., Sirdar-Iskandar, C. (1980). *Les mots du discours*. Paris : Les Editions de Minuit.
- Ducrot, O. (1975). Je trouve que. *Semantikos*, 1(1), 63-88
- Ducrot, O. (1978). Deux mais. *Cahier de linguistique*, 8, 109-120.
- Ducrot, O. (1984). *Le dire et le dit*. Paris : Les Editions de Minuit.
- Ducrot, O. (2001). Quelques raisons de distinguer « locuteurs » et « énonciateurs ». *Polyphonie-linguistique et littéraire*, 3, 19-41.
- Evert, S. (2007). Corpora and collocations. In Lüdeling, A. & Kytö, M. (eds). *Corpus Linguistics : an International Handbook*, II. Mouton de Gruyter, 1212-1248.
- Ferreira, A. (2019). Influence du contexte communicationnel sur l'emploi des constructions syntaxiques de type [V.ØP.] avec les verbes *croire* et *penser*. In *Rencontres des Jeunes Chercheurs en Sciences du Langage 2019*.

- Firth, J. R. (1957). A synopsis of linguistic theory 1930-55. *Studies in linguistic analysis*, Oxford : The Philological Society, 1-32. Reprinted in Palmer (1968). *Selected Papers of J. R. Firth 1952-59*, London : Longmans, 168-205.
- Fraser, B. (1975). Hedged performatives. In Cole, P. & Morgan, J. L. (eds.). *Syntax and Semantics*, vol. 3. New York : Academic Press, 187-210.
- Gachet, F. (2009). Les verbes parenthétiques : un statut syntaxique atypique ? *Linx. Revue des linguistes de l'université Paris X Nanterre*, 61, 13-29.
- Gachet, F. (2015). *Incises de discours rapporté et autres verbes parenthétiques : étude grammaticale*. Paris : Honoré Champion.
- Gadet, F. (1992). *Le français populaire*. Paris : PUF.
- Gauker, C. (1998). What is a context of utterance ? *Philosophical studies*, 91, 149-172.
- Giorgi, A., & Pianesi, F. (2005). Credo (I believe) : Epistemicity and the syntactic representation of the speaker. *Working papers in Linguistics*, 15, 105-152.
- Glikman, J. (2009). Les incises en croire et cuidier en ancien français. *Linx. Revue des linguistes de l'université Paris X Nanterre*, 61, 71-85.
- Goria, E. & Mauri, C. (2018). Il corpus KIParla: una nuova risorsa per lo studio dell'italiano parlato. In Masini, F. & Tamburini, F. (eds.). *CLUB Working Papers in Linguistics*, vol. 2, Bologna: Circolo Linguistico dell'Università di Bologna, 96-116.
- Gosselin, L. (2010). *Les modalités en français - la validation des représentations*. Amsterdam, New York : Rodopi.
- Gosselin, L. (2014). Sémantique des jugements épistémiques : degré de croyance et prise en charge. *Langages*, 1, 63-81.
- Gosselin, L. (2015). L'expression de l'opinion personnelle : « Je crois/pense/trouve/considère/estime que p ». *L'information grammaticale*, 144, 34-40.
- Guerry, M., Catelain, A., Caron, J. (1993). La compréhension de marqueurs modaux : verbes d'attitude propositionnelle et adverbes. *L'Année psychologique*, 93(2), 201-225.
- Guimier, C. (1996). *Les adverbes du français : le cas des adverbes en -ment*. Gap/Paris : Ophrys.
- Guiraud, P. L. (1960). *Problèmes et méthodes de la statistique linguistique*. Presses universitaires de France.

- Halliday, M. A. K. (1985). *An introduction to functional grammar*. Edward Arnold.
- Hare, R. M. (1970). Meaning and speech acts. *Philosophical Review*, 79, 3-24
- Harris, Z. S. (1954). Distributional structure. *Word*, 10(2-3), 146-162.
- Haßler, G. (2014). Étude comparée de l'usage parenthétique des verbes épistémiques dans trois langues romanes. *Discours* [En ligne], 14. URL : <https://doi.org/10.4000/discours.8888>
- Heiden, S., Magué, J.-P., Pincemin, B. (2010). TXM : Une plateforme logicielle open-source pour la textométrie – conception et développement. *JADT 2010 : 10th International Conference on the Statistical Analysis of Textual Data*. Rome, Italie. URL : http://halshs.archives-ouvertes.fr/docs/00/54/97/79/PDF/Heiden_al_jadt2010.pdf
- Hoey, M. (2006). *Lexical priming : a new theory of words and language* [Reprint]. London : Routledge.
- Hooper, J. B. (1975). On assertive predicates. In Kimball, J. P. (ed.). *Syntax and Semantics*, 4, New York : Academic Press, 91-124.
- Infantidou, E. (2001). *Evidentials and relevance*. Amsterdam/Philadelphia : Benjamins.
- Jackendoff, R. (1972). *Semantic interpretation in generative grammar*. MIT Press.
- Jayez, J. & Rossari, C. (2004). Parentheticals as conventional implicatures. In Corblin, F. & Swart, E. de (eds). *Handbook of French Semantics*. Stanford : CSLI Publications, 211-229.
- Jezek, E. (2003). *Classi di verbi tra semantica e sintassi*. Pisa : Edizioni ETS.
- Kerbrat-Orecchioni, C. (1980). *L'énonciation. De la subjectivité dans le langage*. Paris : Armand Colin.
- Kratschmer, A. (2006). Che te ne sembra ? Semantica e pragmatica delle costruzioni italiane con sembrare/parere. *Proceeding : XVI. Skandinaviske Romanistkongres, 24.-27. aug. 2005, Roskilde og København*. Roskilde Universitetscenter, Inst. f. Sprog og Kultur.
- Kronning, H. (1996). *Modalité, cognition et polysémie : sémantique du verbe modal « devoir »*. Uppsala/Stockholm : Almqvist & Wiksell International.
- Lafon, P. (1980). Sur la variabilité de la fréquence des formes dans un corpus. *MOTS*, 1, 127-165.

- Lafon, P. (1981). Analyse lexicométrique et recherche des cooccurrences. *MOTS*, 3, 95-148.
- Lakoff, G. (1973). Hedges : a study in meaning criteria and the logic of fuzzy concepts. *Journal of Philosophical Logic*, 2(4), 458-508.
- Le Querler, N. (1996). *Typologie des modalités*. Caen : Presses Universitaires de Caen.
- Lebart, L. & Salem, A. (1994). *Statistique textuelle*. Paris : Dunod.
- Lo Baido, M. C. (2024). *Forms and Functions of Meta-Discourse. The Case of Comment Clauses in Present-Day Italian*. Berlin : Walter de Gruyter.
- Louw, B. (1993). Irony in the text or insincerity in the writer ? The diagnostic potential of semantic prosodies. In Baker, M., Francis, G., Tognini-Bonelli, E. (eds). *Text and Technology*. Amsterdam : John Benjamins, 157-176.
- Maingueneau, D. (2002). Problèmes d'ethos. *Pratiques : linguistique, littérature, didactique*, 113-114, 55-67.
- Maingueneau, D. (2004). Retour sur une catégorie : le genre. In Adam, J.-M., Grize, J.-B., Ali Bouacha, M. (eds). *Texte et discours : catégories pour l'analyse*. Dijon : Éditions universitaires, 107-118.
- Mauri, C., Ballarè, S., Gorla, E., Cerruti, M., Suriano, F. (2019). KIParla corpus: a new resource for spoken Italian. In Bernardi, R., Navigli, R., Semeraro, G. (eds.). *Proceedings of the 6th Italian Conference on Computational Linguistics CLiC-it*.
- Malrieu, D., & Rastier, F. (2001). Genres et variations morphosyntaxiques. *Traitement Automatique des langues*, 42(2), 548-577.
- Marello, C. (2013). Sembra che e subordinate soggettive. Primi sondaggi in italiano L2 scritto. In Geymonat, F. (ed). *Linguistica applicata con stile. In traccia di Bice Mortara Garavelli*, vol. 15. Edizioni dell'Orso, 79-94.
- Martin, R. (1987). *Langage et croyance : les « univers de croyance » dans la théorie sémantique*. Bruxelles : P. Mardaga.
- Martin, R. (1988). Croire que p/penser que p. *Cahiers d'Études Hispaniques Médiévales*, 7(1), 547-554.
- Martineau, F. (1993). Rection forte et rection faible des verbes : l'ellipse de que en français du Québec et de l'Ontario. *Francophonies d'Amérique*, 3, 79-90.

- Mayaffre, D. (2002). Les corpus réflexifs : entre architextualité et hypertextualité. *Corpus* [En ligne], 1.
- Mayaffre, D., Pincemin, B., Poudat, C. (2019). Explorer, mesurer, contextualiser. Quelques apports de la textométrie à l'analyse de discours. *Langue française*, 203, 101-115.
- Mayaffre, D. (1997). L'Histoire mot à mot. *Cahiers de la Méditerranée*, 54, 159-171.
- Mounin, G. (1974). *Dictionnaire de la linguistique*. Paris : Presses Universitaires de France.
- Muller, C. (1977). *Principes et méthodes de statistique lexicale*. Classiques Hachette.
- Narrog, H. (2012). *Modality, Subjectivity, and Semantic Change : A Cross-Linguistic Perspective*. Oxford : Oxford University Press.
- Nølke, H. (1994). La dilution linguistique des responsabilités : essai de description polyphonique des marqueurs évidentiels « il semble que et il paraît que ». *Langue française*, 102, 84-94.
- Nølke, H., Fløttum, K., Norén, C. (2004). *ScaPoLine. La théorie scandinave de la polyphonie linguistique*. Paris : Editions Kimé.
- Partington, A. (2004). Utterly content in each other's company : Semantic prosody and semantic preference. *International Journal of Corpus Linguistics*, 9(1), 131-156.
- Paveau, M.-A. (2017). *L'analyse du discours numérique. Dictionnaire des formes et des pratiques*. Paris : Hermann.
- 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*, 99, 26-36. URL : https://web.archive.org/web/20190429105402id_/http://icar.univ-lyon2.fr/membres/bpincemin/biblio/pincemin_taln99.pdf
- Pincemin, B. (2011). Sémantique interprétative et textométrie - Version abrégée. *Corpus*, 10, 259-269.
- Pincemin, B. (2022). Sémantique textométrique. In Biglari, A. & Ducard, D.(eds) *La sémantique au pluriel. Théories et méthodes*. Presses universitaires de Rennes, 373-396.
- Pustka, E. (2010). La subordination sans subordonnant en français guadeloupéen - créolisme ou pseudo-créolisme. In Drescher, M. & Neumann-Holzschuh, I. (eds). *La syntaxe de l'oral dans les variétés non-hexagonales du français*. Tübingen : Stauffenburg, 149-166.

- Recanati, F. (1984). Remarques sur les verbes parenthétiques. In Attal, P. & Muller, C. (eds.). *De la syntaxe à la pragmatique : actes du colloque de Rennes, Université de Haute-Bretagne*, vol. 8, John Benjamins, 319-352.
- Rendulić, N. & Kanaan-Caillol, L. (2016). Je crois que, je pense que : valeurs et variation dans un corpus oral diachronique. *SHS Web of Conferences*, 27, 02014. EDP Sciences.
- Ricchiardi, M. & Ježek, E. (2022). Annotating Propositional Attitude Verbs and their Arguments. In *Proceedings of the 18th Joint ACL-ISO Workshop on Interoperable Semantic Annotation within LREC2022*, European Language Resources Association, Marseille, France, 142-149. URL : <https://aclanthology.org/2022.isa-1.19/>
- Riegel, M., Pellat, J.-C., Rioul, R. (2002). *Grammaire méthodique du français* [2^{ème} édition]. Paris : Presses Universitaires de France.
- Ross, J. R. (1973). Slifting. In Ross, J. R., Gross, M., Halle, M., Schützenberger, M. P. (eds). *The Formal Analysis of Natural Language : Proceedings of the First International Conference*. The Hague : Mouton, 133-170.
- Rossari, C., Montrichard, C., Pamuksaç, A., Ricci, C., Sanvido, L. (2022). 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., Sanvido, L. (2024). Vers un schéma cognitif de l'exemplification. Étude textométrique et contrastive de formes de l'exemplification en français et en italien. *Langue française*, 222(2), 59-76.
- Rouanne, L. (2016). Modalisation et expression de la réserve : si on peut dire, si je puis dire, si j'ose dire. In Rouanne, L. & Anscombre, J.-C. (eds). *Histoires de dire. Petit glossaire des marqueurs formés sur le verbe dire*. Bern : Peter Lang, 271-291.
- Roulet, E. (1987). Complétude interactive et connecteurs reformulatifs. *Cahiers de linguistique française*, 8, 111-140.
- Roulet, E., Auchlin, A., Moeschler, J., Rubattel, C., Schelling, M. (1987). *L'articulation du discours en français contemporain* [2^e édition]. Bern : Peter Lang.
- Roventini, A., Alonge, A., Bertagna, F., Calzolari, N., Zampolli, A., Girardi, C., Magnini, B., Marinelli, R., Speranza, M., Cancila, J. (2003). Italwordnet : Building a large semantic database for the automatic treatment of the italian language. In Zampolli, A., Calzolari,

- N., Cignoni, L. (eds.). *Computational Linguistics in Pisa, special issue*, vol. 2, Pisa/Roma : Istituti editoriali e poligrafici internazionali, 745-791.
- Rundell, M. (2018). Searching for extended units of meaning-and what to do when you find them. *Lexicography*, 5(1), 5-21.
- Russell, B. (1985). *The Philosophy of Logical Atomism*. Routledge.
- Sabatini, F. (1985). L'italiano dell'uso medio : una realtà tra le varietà linguistiche italiane. In Holtus, G. & Radtke, E. (eds.). *Gesprochenes italienisch in Geschichte und Gegenwart*, Tübingen : Narr, 154-184.
- Sanvido, L. (2024). *Je crois vraiment, je dis simplement : étude statistique et comparative français-italien des associations entre adverbes et verbes d'opinion*. Cahiers de praxématique [En ligne], 82. URL : <https://doi.org/10.4000/12evx>
- Savoia, L. M. (2008). Variazione e mescolanza linguistica nei sistemi arbëreshë : codemixing, prestiti e convergenza in condizioni di bilinguismo. In Savoia, L. M. (ed.). *Studi sulle varietà arbëreshe*. Cosenza : Università della Calabria, 1-62.
- Schena, L., Prandi, M., Mazzoleni, M. (2000). *Intorno al congiuntivo*. Bologna : Clueb.
- Schneider, S. (1999). *Il congiuntivo tra modalità e subordinazione : uno studio sull'italiano parlato*. Roma : Carocci.
- Schneider, S. (2007). *Reduced Parenthetical Clauses as Mitigators : a Corpus Study of Spoken French, Italian and Spanish*. Amsterdam : John Benjamins.
- Schneider, S. (2009). Les clauses parenthétiques dans les textes de la Nouvelle-France du 17e et du 18e siècle : aspects pragmatiques, syntaxiques et diachroniques. *Linx. Revue des linguistes de l'université Paris X Nanterre*, 61, 87-102.
- Searle, J. R. (1982). *Sens et expression. Études de théorie des actes du langage*. Paris : Les Editions de Minuit.
- Simon-Vandenberg, A. M. (2000). The functions of I think in political discourse. *International Journal of Applied Linguistics*, 10(1), 41-63.
- Sinclair, J. (1991). *Corpus, Concordance, Collocation*. Oxford University Press.
- Sinclair, J. (1996). The search for units of meaning. *Textus*, 9(1), 75-106.
- Sinclair, J. (1999). The lexical item. In Weigand, E. (ed). *Contrastive Lexical Semantics*. Amsterdam : John Benjamins.

- Sitri, F. & Barats, C. (2017). Constituer un corpus en analyse du discours, un moment crucial. In Née, E. (ed.). *Méthodes et outils informatiques pour l'analyse des discours*. Rennes : Presses universitaires de Rennes, 41-62.
- Sitri, F. & Tutin, A. (2016). Présentation, *Lidil*, 53, 5-18.
- Soliman, L. T. (2010). Bivalence modale dans les complétives régies par des verbes épistémiques : perspective contrastive (italien/français). *Modèles linguistiques*, 62, 63-81.
- Stalnaker, R. (2002). Common ground. *Linguistics and philosophy*, 25, 701-721.
- Stewart, D. (2010). *Semantic prosody a critical evaluation*. New York : Routledge.
- Strik Lievers, F. (2012). *Sembra ma non è. Studio semantico-lessicale sui verbi con complemento predicativo*. Firenze : Accademia della Crusca.
- Thompson, S. A. & Mulac, A. (1991). The discourse conditions for the use of the complementizer that in conversational English. *Journal of pragmatics*, 15(3), 237-251.
- Tognini-Bonelli, E. (2001). *Corpus linguistics at work*. Amsterdam : J. Benjamins.
- Traugott, E. C. (1995). Subjectification in grammaticalization. In Stein, D. & Wright, S. (eds.). *Subjectivity and subjectivisation : Linguistic perspectives*. Cambridge University Press, 31-54.
- Tuchais, S. (2012). Les verbes d'opinion et la variabilité des jugements : le cas de *je considère que p* et *j'estime que p*. *Bulletin d'études de linguistique française*, 46, 35-50.
- Tuchais, S. (2014). *Comment dire ce que «je» pense en japonais et en français : étude contrastive de l'expression de l'opinion personnelle*. Thèse de doctorat, Paris, EHESS.
- Tuchais, S. (2020). Les modulateurs d'assertion en français et en japonais du point de vue de la constructionnalisation. *Langages*, 220(4), 43-64.
- Urmson, J. O. (1952). Parenthetical verbs. *Mind*, 61(244), 480-496.
- Ursini, F. (2007). L'italiano dei Presidenti e l'italiano degli Italiani. In Cortellazzo, M. A. & Tuzzi, A. *Messaggi dal Colle. I discorsi di fine anno dei presidenti della Repubblica*. Venezia : Marsilio, 195-206.
- Venier, F. (1991). *La modalizzazione assertiva. Avverbi modali e verbi parentetici*. Milano : Franco Angeli.

- Vigier, D. (2017). « *Autour* » des prépositions *en, dans, dedans*. Vers une approche diachronique sur corpus outillé. Habilitation à diriger des recherches, Université Toulouse Jean Jaurès.
- Wiesmath, R. (2002). Présence et absence du relatif et conjonctif *que* dans le français acadien : tendances contradictoires ? In Pusch, C. & Raible W. (eds). *Romanistische Korpuslinguistik – Korpora und gesprochene Sprache/Romance Corpus Linguistics – Corpora and Spoken Language*, Tübingen : Gunter Narr Verlag, 393-408.
- Willems, D. & Blanche-Benveniste, C. (2010). Verbes « faibles » et verbes à valeur épistémique en français parlé : *il me semble, il paraît, j’ai l’impression, on dirait, je dirais*. In Iliescu, M., Danler, P., Siller-Runggaldier, H. (eds.). *Actes du XXVe Congrès International de Linguistique et de Philologie Romanes*, Tome IV. Berlin : Mouton de Gruyter, 565-576.
- Willems, D. (1981). *Syntaxe, lexicque et sémantique : les constructions verbales*. Gent : Rijksuniversiteitte Gent.
- Zufferey, S. (2020). *Introduction à la linguistique de corpus*. London : ISTE Editions Ltd.

8.2 Dictionnaires, encyclopédies et traducteurs en ligne

- « Accusativo preposizionale ». In *Treccani.it – Enciclopedia dell’Italiano (2010)*. Roma : Istituto della Enciclopedia Italiana. URL : [https://www.treccani.it/enciclopedia/accusativo-preposizionale_\(Enciclopedia-dell%27Italiano\)/](https://www.treccani.it/enciclopedia/accusativo-preposizionale_(Enciclopedia-dell%27Italiano)/) (consulté le 07.02.2023)
- « Considérer ». In *Dictionnaire en ligne Larousse*. Paris : Editions Larousse. URL : <https://www.larousse.fr/dictionnaires/francais/consid%C3%A9rer/18387> (consulté le 05.12.2024)
- « Considérer ». In *TLFi - Trésor de la langue Française informatisé*. ATILF - CNRS & Université de Lorraine. URL : <https://www.cnrtl.fr/definition/consid%C3%A9rer> (consulté le 05.12.2024)
- « Dichiarativo ». In *Treccani.it - Vocabolario online*. Roma : Istituto della Enciclopedia Italiana. URL :

<https://www.treccani.it/vocabolario/dichiarativo/#:~:text=%5Bdal%20lat.,pi%C3%B9%20spesso%20col%20nome%20lat> (consulté le 07.02.2023)

« Ella ». In *Treccani.it - Vocabolario online*. Roma : Istituto della Enciclopedia Italiana. URL : <https://www.treccani.it/vocabolario/ella/> (consulté le 07.02.2023)

« Ritenére ». In *Treccani.it - Vocabolario online*. Roma : Istituto della Enciclopedia Italiana. URL : <https://www.treccani.it/vocabolario/ritenere/> (consulté le 05.12.2024)

DeepL Traduction. URL : <https://www.deepl.com/fr/translator>

Dictionnaire bilingue français-italien en ligne Larousse. Paris : Editions Larousse. URL : <https://www.larousse.fr/dictionnaires/francais-italien>

Dictionnaire bilingue français-italien en ligne Sansoni. Milano : Edigeo s.r.l. URL : https://dizionari.corriere.it/dizionario_francese/

Il nuovo vocabolario di base della lingua italiana. Tullio De Mauro, Internazionale. URL : <https://dizionario.internazionale.it>

Les verbes français. Dictionnaire électronique des mots. URL : http://rali.iro.umontreal.ca/LVF_DEM/LVF-DEM-req.html

Reverso Context. URL: <https://context.reverso.net/traduzione/francese-italiano/>

Treccani.it - Enciclopedia online. Roma : Istituto della Enciclopedia Italiana. URL : <https://www.treccani.it/enciclopedia/>

Treccani.it - Vocabolario online. Roma : Istituto della Enciclopedia Italiana. URL : <https://www.treccani.it/vocabolario/>

8.3 Corpus et bases de données

Corpus *CRIPCO* (*Cross-document Italian People Coreference*) : <http://metashare.elda.org/repository/browse/cripco/4ab6b4560f8d11e2b7c50015c5ed7672044ca04410a24fe4b0b894bcaa77075d/>

Corpus *Discours présidentiels français de 1958 à aujourd'hui*. In *Hyperbase Web Edition*, Université Côte d'Azur. URL : <http://hyperbase.unice.fr/hyperbase/>

Corpus *OPUS*. In *Sketch Engine*. URL : <https://www.sketchengine.eu/>

Europresse. URL : <https://www.europresse.com/>

Ortolang. Plate-forme d'outils et de ressources linguistiques pour un traitement optimisé de la langue française URL : www.ortolang.fr

8.4 Logiciels et outils

Notepad++. URL : <https://notepad-plus-plus.org>

R Studio. URL : <https://posit.co/download/rstudio-desktop/>

Textométrie // TXM. URL : <https://txm.gitpages.huma-num.fr/textometrie/en/index.html>

Treetagger, A part-of-speech tagger for many languages. LMU München. URL : <https://www.cis.uni-muenchen.de/~schmid/tools/TreeTagger/>

Web Scraper. URL : <https://webscraper.io/>

9. Annexes

9.1 Paradigmes des formes

Formes françaises exclues à cause du critère de fréquence uniquement	Formes françaises exclues à cause du test de <i>à mon avis</i> uniquement	Formes françaises exclues à cause des deux critères
j'ajoute que j'ai l'impression que j'ai la conviction que j'ai le sentiment que j'avoue que je doute que je suis persuadé j'imagine que je précise que je répète que	j'accepte que je comprends que je constate que je crains que je demande que je me demande si j'espère que je regrette que je sais que je souhaite que je vois que je veux que	j'aperçois que je m'aperçois que j'attends que j'ai conscience que j'ai peur que je confirme que j'entends que je m'étonne que je suis conscient que je suis heureux que j'ignore que je note que j'observe que j'oublie que je me rappelle que je rappelle que je reconnais que je me réjouis que je remarque que je sens que je me souviens que

Tableau 7 : verbes d'indication de point de vue français exclus sur la base des critères de fréquence et de remplacement par « à mon avis ».

Formes italiennes exclues à cause du critère de fréquence uniquement	Formes italiennes exclues à cause du test de <i>secondo me</i> uniquement	Formes italiennes exclues à cause des deux critères
aggiungo che assicuro che ho l'impressione che dubito che immagino che rispondo che suppongo che mi sembra che trovo che	mi auguro che auspico che ricordo che so che spero che	mi accorgo che auguro che ho fiducia che ho la speranza che capisco che confido che sono consapevole che sono contento che sono fiducioso che osservo che mi ricordo che sento che temo che vedo che

Tableau 8 : verbes d'indication de point de vue italiens exclus sur la base des critères de fréquence et de remplacement par « secondo me ».

9.2 Classes sémantiques

CS de la classe <i>Construction argumentative</i>	CS de la classe <i>Prise en charge</i>	CS de la classe <i>Modalité</i>	CS de la classe <i>Evaluation</i>
donc mais parce que	moi personnellement	devoir falloir nécessaire possible pouvoir probablement	assez beaucoup bon brave davantage exagérément fondamental important intelligent mauvais noble normal peu préférable primordial raisonnable sage surtout tactique très urgent utile véridique vulnérable

Tableau 9 : composition des classes sémantiques du niveau intermédiaire (français).

CS de la classe <i>Construction</i> <i>argumentative</i>	CS de la classe <i>Prise en charge</i>	CS de la classe <i>Modalité</i>	CS de la classe <i>Evaluation</i>
anzi ma perché però	francamente io onestamente personalmente	bisognare bisogno dovere (n.) dovere (v.) doveroso forse necessario occorrere possibilità potere	accettabile antipatico bello ben bene benissimo buon buono certo competitivo così davvero deleterio difficile duraturo eccessivo fondamentale forte fortemente già giusto grande importante maturo meglio migliore molto normale opportuno ottimo più ragionevolmente rigoroso seriamente sicuramente sicuro simile soddisfatto sufficiente tanto troppo veramente

Tableau 10 : composition des classes sémantiques du niveau intermédiaire (italien).

9.3 Listes exhaustives de tous les cooccurrents spécifiques

Nous présentons ci-dessous les listes exhaustives de tous les cooccurrents spécifiques des verbes d'indication de point de vue (français et italiens). Les cooccurrents spécifiques qui ont été retenus pour l'établissement du profil combinatoire sont ceux qui sont colorés selon leur classe sémantique d'appartenance (*Construction argumentative* en violet, *Prise en charge* en jaune, *Modalité* en vert et *Evaluation* en bleu).

Attention : dans ces listes exhaustives, les cooccurrents qui appartiennent à la catégorie morphosyntaxique « verbe » sont affichés séparément selon le temps verbal. Ainsi, il est possible que le verbe *pouvoir* apparaisse plusieurs fois : par exemple, une fois pour signaler qu'il est conjugué au conditionnel, une fois pour signaler qu'il est conjugué au présent de l'indicatif (voir l'exemple de *je pense que* dans PRES-FR). Etant donné que cette distinction de temps verbal n'est pas approfondie dans le cadre de l'étude des verbes retenus pour la classe *Modalité*, ceux-ci ont été traités en tant que lemmes dans les tableaux des niveaux micro (Figure 14, Figure 15, Figure 20 et Figure 21). Par conséquent, les indices de spécificité entre verbes affichés par temps verbal et par lemme peuvent varier.

Je crois que

Liste des cooccurrents spécifiques de *je crois que* dans l'empan gauche de 4 items (tableau de gauche) et dans l'empan droit de 10 items (tableau de droite) dans le corpus PRES-FR.

Occ	Fréq	CoFréq	Indice	DistMoy
SENT_.	101711	639	203,72	0,57
KON_parce	5095	47	21,46	1,09
KON_mais	13735	67	15,67	0,27
KON_et	55986	160	13,62	0,69
PUN_	184896	396	12,26	0,61
INT_oui	1619	17	9,09	1,29
SENT_?	5858	29	7,37	0,59
PUN_...	1979	15	6,30	0,13
INT_eh	721	9	5,73	1,67
ADV_alors	2999	17	5,35	1,06
PRO:PER_moi	3371	17	4,71	0,59
NOM_penser	40	3	4,51	0,67
PRO:PER_le	7492	25	3,64	2,16
ADV_absolument	416	5	3,36	1,60
NOM_contexte	110	3	3,20	1,33
NOM_or	467	5	3,13	0,60
NOM_radio	119	3	3,10	0,67

Occ	Fréq	CoFréq	Indice	DistMoy
VER:pres_falloir	6540	120	43,69	2,35
PRO:PER_il	38834	338	43,64	2,06
VER:pres_être	55314	407	36,10	3,19
PRO:DEM_ce	64560	414	24,94	2,75
VER:pres_pouvoir	6426	77	17,53	2,61
PRO:PER_on	17947	120	8,76	2,12
NAM_Français	3564	39	8,24	3,23
ADJ_bon	1892	26	7,59	4,15
ADV_très	4635	44	7,41	4,30
DET:ART_un	50127	262	7,32	4,23
PRO:DEM_cela	10001	72	6,70	2,93
ADJ_important	1721	23	6,60	3,35
ADJ_nécessaire	1242	18	5,80	3,83
ADV_exagérément	27	4	5,51	4,75
NUM_deux	2673	27	5,31	3,41
ADJ_véridique	2	2	4,86	1,00
VER:pres_devoir	5843	44	4,84	3,66
PRO:IND_tout	12003	75	4,79	3,16
VER:ppe_exagérer	14	3	4,75	2,00
NOM_fait	2133	22	4,59	5,09
VER:pres_valoir	233	7	4,51	1,86
VER:infi_dire	4958	38	4,43	4,21
NOM_idée	1063	14	4,22	3,79
NOM_tactique	22	3	4,13	1,33
VER:pres_avoir	47298	227	4,11	3,77
NOM_point	1693	18	4,04	4,06
ADV_tout	4959	36	3,78	4,11
VER:infi_réfléchir	143	5	3,68	3,40
NOM_article	145	5	3,65	3,80
NOM_vue	791	11	3,64	4,82
VER:pres_correspondre	151	5	3,57	1,20
NOM_trilinguisme	7	2	3,55	3,50
NOM_erreur	238	6	3,53	2,83
ADV_surtout	687	10	3,52	4,60
VER:pres_confondre	38	3	3,41	2,00
ADV_peu	2305	20	3,27	5,10
ADJ_français	3419	26	3,20	5,46
NOM_tort	105	4	3,18	3,50
NOM_démocratie	905	11	3,16	6,27
VER:infi_distinguer	48	3	3,11	2,67
NOM_rôle	922	11	3,09	4,45
NOM_caractéristique	49	3	3,09	3,33

Liste des cooccurents spécifiques de *je crois que* dans l'empan gauche de 4 items (tableau de gauche) et dans l'empan droit de 10 items (tableau de droite) dans le corpus LM.

Occ	Fréq	CoFréq	Indice	DistMoy
SENT_.	668464	137	37,68	0,74
SENT_?	26671	28	25,41	0,61
PUN:cit_«	157502	47	19,07	0,06
KON_mais	46185	20	11,28	0,05
ABR_\	1908	6	8,74	0,83
NAM_Dolto	11	2	6,77	1,00
DET:POS_mon	6928	7	6,74	2,00
PUN_...	16883	9	6,17	0,67
SENT_!	9408	6	4,73	0,33
NOM_affirmation	237	2	4,06	0,50
NOM_chapitre	444	2	3,52	0,50
PRO:PER_moi	2299	3	3,50	0,67
PUN_:	58873	11	3,22	0,91
ADV_clairement	645	2	3,20	0,50
NOM_façon	3311	3	3,04	1,33

Occ	Fréq	CoFréq	Indice	DistMoy
PRO:PER_il	146572	80	23,04	2,74
PRO:PER_je	26613	33	19,76	2,85
VER:pres_être	194652	86	18,98	3,47
VER:pres_falloir	8549	19	16,15	2,11
PRO:DEM_ce	159959	68	14,37	2,69
PRO:PER_nous	25520	21	9,61	1,90
ADV_ne	117348	45	8,42	2,71
ADV_personne	784	5	6,93	0,00
VER:pres_avoir	202806	60	6,91	3,48
PRO:DEM_cela	12324	12	6,55	1,42
NOM_gens	3208	7	6,32	2,86
VER:pres_devoir	11048	11	6,15	3,55
DET:POS_mon	6928	9	6,07	4,00
NOM_chose	4680	7	5,25	4,00
VER:pres_sous-estimer	33	2	4,99	0,50
PRO:IND_tout	30801	16	4,94	3,56
VER:cond_avoir	7375	8	4,90	2,50
VER:impf_accorder	46	2	4,70	4,50
ADV_beaucoup	6062	7	4,53	3,14
ADJ_brave	59	2	4,48	1,00
NOM_tort	461	3	4,38	3,00
PRO:PER_y	27338	14	4,34	3,57
NOM_retournement	93	2	4,08	3,00
NOM_diplomatie	605	3	4,03	3,33
ADJ_possible	3155	5	4,02	3,00
PRO:PER_vous	9937	8	4,00	2,50
VER:pres_pouvoir	15938	10	3,96	3,10
PRO:PER_on	37542	16	3,94	2,63
VER:futu_changer	111	2	3,93	1,50
ADV_là	7877	7	3,83	4,29
PRO:PER_me	8020	7	3,78	3,43
PRO:REL_que	42943	17	3,77	5,59
ADV_fermeement	137	2	3,75	2,00
ADJ_tactique	157	2	3,63	3,00
ADV_pas	80157	25	3,61	4,28
NOM_politologue	176	2	3,53	4,50
VER:infi_surmonter	215	2	3,36	2,00
VER:ppe_évoluer	227	2	3,31	2,00
NAM_Colombie_	248	2	3,23	3,00
VER:infi_ignorer	254	2	3,21	2,50
ADJ_vulnérable	259	2	3,20	4,00
NOM_écrit	299	2	3,07	2,50
VER:ppe_réussir	1363	3	3,00	2,33

Je pense que

Liste des cooccurrents spécifiques de *je pense que* dans l'empan gauche de 4 items (tableau de gauche) et dans l'empan droit de 10 items (tableau de droite) dans le corpus PRES-FR :

Occ	Fréq	CoFréq	Indice	DistMoy
SENT .	101711	480	135,76	0,61
PRO:PER_moi	3371	33	18,35	0,85
KON_mais	13735	64	17,82	0,39
ADV_personnellement	160	11	15,58	0,64
KON_parce	5095	33	13,29	1,00
PUN_	184896	335	11,39	0,58
KON_et	55986	128	10,03	0,62
ADV_pourquoi	1953	15	7,38	0,00
ADV_donc	5035	23	6,76	0,74
SENT !	3240	18	6,63	0,06
ADV_non	3628	19	6,57	1,16
INT_eh	721	8	5,37	1,38
ADV_bien	6690	23	4,75	1,39
VER:infi_féliciter	43	3	4,65	1,67
VER:pper_fiscaliser	6	2	4,64	0,50
ADV_auparavant	47	3	4,53	0,67
NOM_magistrature	51	3	4,43	1,33
ADV_voilà	1011	8	4,32	1,38
PUN_...	1979	11	4,31	0,91
NOM_référendum	328	5	4,20	1,00
SENT_?	5858	20	4,18	0,75
INT_oui	1619	9	3,63	1,11
PUN_)	340	4	3,04	0,75
NOM_cour	151	3	3,03	2,00
VER:pper_répondre	151	3	3,03	1,33

Occ	Fréq	CoFréq	Indice	DistMoy
PRO:PER_il	38834	272	33,17	1,97
VER:pres_falloir	6540	66	15,44	1,98
PRO:DEM_ce	64560	316	14,48	3,06
ADJ_bon	1892	31	12,75	3,97
NOM_erreur	238	12	10,72	4,17
VER:pres_devoir	5843	52	10,39	2,62
VER:pres_être	55314	254	9,19	3,34
VER:cond_être	2004	25	8,07	2,88
VER:infi_être	6766	51	7,85	4,25
ADJ_important	1721	21	6,74	4,57
ADJ_sage	45	5	6,52	4,60
VER:cond_falloir	290	9	6,41	1,00
DET:ART_un	50127	217	6,14	4,06
ADJ_utile	386	9	5,39	2,78
VER:cond_pouvoir	1079	14	5,01	4,79
VER:pres_pouvoir	6426	41	4,75	3,29
VER:pres_penser	2213	20	4,55	6,05
ADV_assez	776	11	4,41	4,09
VER:pres_avoir	47298	195	4,39	3,69
NOM_chose	2928	23	4,20	5,09
VER:futu_être	3745	27	4,19	2,74
NOM_sujet	1296	14	4,16	4,93
NAM_Français	3564	26	4,14	3,00
VER:cond_devoir	492	8	3,77	2,38
ADV_très	4635	30	3,77	4,23
VER:cond_avoir	1422	14	3,75	3,50
ADJraisonnable	189	5	3,49	4,60
NOM_tort	105	4	3,48	3,25
ADJ_urgent	44	3	3,46	3,33
ADVdavantage	701	9	3,41	3,89
ADJ_sain	48	3	3,35	1,67
PRO:PER_y	15141	71	3,33	3,48
NOM_raison	1756	15	3,32	4,33
PRO:PER_on	17947	81	3,22	2,51
ADJ_intelligent	55	3	3,17	2,00
ADJ_démocrate	13	2	3,14	4,50
NOM_oxygène	13	2	3,14	2,50
VER:infi_faire	6503	36	3,12	4,47
NOM_assiette	15	2	3,02	4,50

Liste des cooccurrents spécifiques de *je pense que* dans l'empan gauche de 4 items (tableau de gauche) et dans l'empan droit de 10 items (tableau de droite) dans le corpus LM :

Occ	Fréq	CoFréq	Indice	DistMoy
PUN:cit_«	157502	89	45,27	0,16
SENT_.	668464	174	41,64	0,74
SENT_?	26671	36	31,09	0,36
PUN_:	58873	34	17,95	1,03
KON_mais	46185	24	11,94	0,00
ADV_personnellement	184	3	6,31	0,67
NAM_Médiapart	18	2	6,03	1,50
PRO:PER_moi	2299	5	5,87	0,60
DET:POS_mon	6928	7	5,77	2,14
NAM_Quand	659	3	4,66	0,00
NOM_ban	91	2	4,60	1,50
VER:pres_convaincre	106	2	4,47	1,00
ADJ_naïf	155	2	4,14	1,00
NOM_commentateur	188	2	3,97	1,50
NOM_soin	1235	3	3,85	2,00
KON_car	6292	5	3,79	0,00
NOM_grève	1620	3	3,51	1,67
ADJ_optimiste	387	2	3,35	1,00
VER:pres_avancer	421	2	3,27	0,50
NOM_télé	456	2	3,21	0,50
PRO_qui	481	2	3,16	1,50

Occ	Fréq	CoFréq	Indice	DistMoy
PRO:PER_il	146572	105	27,60	1,85
PRO:PER_je	26613	43	24,23	4,12
PRO:PER_nous	25520	42	23,99	2,17
VER:pres_aller	11492	26	18,41	2,15
VER:pres_falloir	8549	23	17,98	2,00
ADV_ne	117348	70	14,70	3,64
PRO:PER_on	37542	36	13,51	2,94
PRO:DEM_cela	12324	20	11,74	1,55
PRO:DEM_ce	159959	77	11,46	2,74
VER:pres_être	194652	86	10,83	3,49
VER:pres_pouvoir	15938	21	10,63	2,10
NOM_gens	3208	10	8,83	4,70
VER:subi_M	1295	7	7,97	0,00
PRO:PER_y	27338	23	7,88	2,78
ADV_pas	80157	42	7,57	4,36
VER:pres_avoir	202806	76	6,75	3,64
VER:infi_dire	4158	9	6,68	5,44
ADV_assez	2287	7	6,32	4,14
ADJ_bon	7679	11	6,25	4,18
VER:cond_devoir	6818	10	5,83	2,20
KON_si	16689	15	5,72	5,60
VER:futu_avoir	4199	8	5,60	1,63
VER:cond_avoir	7375	10	5,53	3,20
VER:infi_être	23175	17	5,23	3,24
NOM_épaule	285	3	4,56	3,33
KON_parce	4304	7	4,54	3,86
NAM_Ferrand	45	2	4,42	1,00
DET:POS_mon	6928	8	4,06	0,50
VER:infi_porter	1168	4	4,02	4,50
VER:futu_voter	78	2	3,94	4,50
VER:infi_changer	1259	4	3,90	3,50
VER:pres_espérer	1307	4	3,84	4,75
ADJ_préférable	89	2	3,83	1,00
NOM_erreur	1330	4	3,81	1,75
NOM_sursaut	92	2	3,80	2,00
ADJ_utile	541	3	3,74	5,00
NOM_chantage	103	2	3,70	2,50
VER:pres_sommer êtr	2656	5	3,68	2,00
ADJ_noble	124	2	3,54	2,00
VER:infi_garder	633	3	3,54	2,67
VER:pper_peser	139	2	3,44	2,00
NOM_chose	4680	6	3,42	4,67
VER:pres_devoir	11048	9	3,38	3,33
ADV_forcément	739	3	3,34	5,67
ADV_probablement	747	3	3,33	4,00
NOM_volonté	1863	4	3,26	4,00
VER:infi_bouger	179	2	3,22	2,00
NOM_convergence	184	2	3,20	3,00
PRO:PER_la le	25374	14	3,18	2,36
VER:pres_agir	3463	5	3,16	3,00
VER:infi_décrire	194	2	3,15	1,50
VER:cond_être	7419	7	3,11	2,29
VER:cond_falloir	920	3	3,07	0,67
VER:pres_estimer	3654	5	3,06	5,60
VER:infi_faire	17588	11	3,05	2,73
PRO:PER_elle	45473	20	3,03	2,85
ADV_ensemble	971	3	3,00	3,33

Je considère que

Liste des cooccurrents spécifiques de *je considère que* dans l'empan gauche de 4 items (tableau de gauche) et dans l'empan droit de 10 items (tableau de droite) dans le corpus PRES-FR :

Occ	Fréq	CoFréq	Indice	DistMoy
SENT_.	101711	105	28,72	0,78
KON_parce	5095	26	23,21	1,00
PRO:PER_moi	3371	17	15,34	0,65
ADJ_parlementaire	285	4	5,79	0,75
KON_mais	13735	15	4,85	0,07
PUN_	184896	80	3,82	0,81
KON_et	55986	32	3,67	0,63
NOM_or	467	3	3,46	1,00
KON_car	1445	4	3,07	0,00
ADV_personnellement	160	2	3,00	0,50

Occ	Fréq	CoFréq	Indice	DistMoy
VER:pres_être	55314	98	15,06	3,30
NOM_devoir	743	10	9,59	2,60
PRO:DEM_ce	64560	81	5,95	2,40
VER:pres_devoir	5843	16	5,15	2,31
ADJ_fondamental	364	4	3,83	3,25
ADJ_primordial	32	2	3,61	1,50
DET:ART_un	50127	58	3,55	4,17
NOM_fonction	807	5	3,51	3,60
VER:pres_régler	39	2	3,44	2,00
NOM_apport	40	2	3,42	1,50
VER:infi_frapper	41	2	3,39	2,50
NOM_orientation	207	3	3,33	0,67
NAM_France	9602	17	3,17	2,76
ADJ_normal	548	4	3,16	2,25
ADJ_français	3419	9	3,05	4,89
NOM_pluralisme	64	2	3,01	0,50

Liste des cooccurrents spécifiques de *je considère que* dans l'empan gauche de 4 items (tableau de gauche) et dans l'empan droit de 10 items (tableau de droite) dans le corpus LM :

Occ	Fréq	CoFréq	Indice	DistMoy
KON_mais	46185	4	3,54	0,00
KON_parce	4304	2	3,39	0,50
SENT_.	668464	13	3,24	0,54

Occ	Fréq	CoFréq	Indice	DistMoy
DET:POS_mon	6928	5	6,81	1,20
VER:pres_être	194652	13	4,49	4,62
ADV_pas	80157	8	4,12	4,50
NOM_recul	911	2	3,94	3,00
NOM_chance	1683	2	3,41	2,50
ADJ_historique	2106	2	3,22	1,00
VER:pres_devoir	11048	3	3,04	4,33

Je trouve que

Liste des cooccurrents spécifiques de *je trouve que* dans l'empan gauche de 4 items (tableau de gauche) et dans l'empan droit de 10 items (tableau de droite) dans le corpus PRES-FR :

Occ	Fréq	CoFréq	Indice	DistMoy
SENT_.	101711	44	11,65	0,82
PRO:PER_moi	3371	9	9,10	0,89
PUN_...	1979	4	3,88	0,00
ADV_personnellement	160	2	3,70	0,50
ADV_ailleurs	2360	4	3,59	0,75
PUN_	184896	41	3,41	0,41
SENT !	3240	4	3,08	1,50
NOM_réflexion	351	2	3,02	1,00

Occ	Fréq	CoFréq	Indice	DistMoy
VER:pres_être	55314	43	6,96	2,23
ADV_très	4635	11	6,41	3,91
PRO:DEM_ce	64560	45	5,98	2,18
VER:pres_ressembler	51	2	3,91	2,50
PRO:PER_il	38834	27	3,82	2,22
DET:ART_un	50127	31	3,42	3,42
PUN_...	1979	5	3,34	4,20
NOM_tort	105	2	3,28	3,00

Liste des cooccurrents spécifiques de *je trouve que* dans l'empan gauche de 4 items (tableau de gauche) et dans l'empan droit de 10 items (tableau de droite) dans le corpus LM :

Occ	Fréq	CoFréq	Indice	DistMoy
SENT_?	26671	6	6,85	0,33
SENT_.	668464	19	5,34	1,00
PUN:cit_«	157502	9	5,02	0,22
ADV_largement	1327	2	4,20	0,50
PRO:PER_moi	2299	2	3,73	0,50
PUN:cit_"	96195	6	3,72	1,83

Occ	Fréq	CoFréq	Indice	DistMoy
VER:infi_déchoir déchoir	41	2	6,43	4,00
ADJ_ironique	142	2	5,35	2,00
VER:pres_être	194652	16	5,21	2,69
VER:ppre_prévoir	167	2	5,21	3,00
NOM_remarque	371	2	4,51	0,50
PRO:DEM_ce	159959	13	4,30	1,54
ADJ_physique	835	2	3,81	3,50
ADJ_critique	862	2	3,78	1,00
PRO:PER_on	37542	6	3,76	1,33
PRO:PER_il	146572	11	3,43	1,45
ADJ_mauvais	2058	2	3,03	2,00

Je dis que

Liste des cooccurrents spécifiques de *je dis que* dans l'empan gauche de 4 items (tableau de gauche) et dans l'empan droit de 10 items (tableau de droite) dans le corpus PRES-FR :

Occ	Fréq	CoFréq	Indice	DistMoy
SENT_.	101711	47	11,47	0,72
KON_mais	13735	13	6,82	0,38
KON_quand	2625	6	5,56	0,00
ADV_là	4411	7	5,35	0,57
ADV_alors	2999	6	5,23	0,83
PRO:PER_moi	3371	5	3,84	0,60
VER:pres_développer	144	2	3,68	1,00
VER:infi_augmenter	217	2	3,33	1,50

Occ	Fréq	CoFréq	Indice	DistMoy
VER:pres_falloir	6540	17	9,26	2,59
PRO:PER_il	38834	39	7,70	1,85
ADV_ne	35612	29	4,23	2,62
NOM_régulation	59	2	3,66	3,00
ADJ_mauvais	333	3	3,59	4,00
NOM_alternance	122	2	3,04	0,50

Liste des cooccurrents spécifiques de *je dis que* dans l'empan gauche de 4 items (tableau de gauche) et dans l'empan droit de 10 items (tableau de droite) dans le corpus LM :

Occ	Fréq	CoFréq	Indice	DistMoy
NOM_priorité	1229	2	4,60	0,50
PRO:PER_moi	2299	2	4,06	0,50
SENT_.	668464	13	3,85	1,31
KON_quand	6788	2	3,12	0,00
DET:POS_mon	6928	2	3,11	1,50

Occ	Fréq	CoFréq	Indice	DistMoy
VER:pres_falloir	8549	4	5,06	3,00
PRO:DEM_cela	12324	4	4,44	2,00
PRO:PER_il	146572	10	4,16	2,00
VER:infi_améliorer	882	2	4,09	1,00
PRO:PER_la le	25374	4	3,25	3,00
PRO:PER_je	26613	4	3,17	2,00

Il me semble que

Liste des cooccurrents spécifiques de *il me semble que* dans l'empan gauche de 4 items (tableau de gauche) et dans l'empan droit de 10 items (tableau de droite) dans le corpus PRES-FR :

Occ	Fréq	CoFréq	Indice	DistMoy
SENT_.	101711	27	7,50	0,30
NOM_échéance	138	2	4,25	1,00

Occ	Fréq	CoFréq	Indice	DistMoy
VER:cond_devoir	492	5	7,24	3,00
NOM_question	2628	6	4,82	5,17
NOM_moment	1908	5	4,39	4,40
NOM_initiative	689	3	3,45	2,67
NOM_phénomène	171	2	3,27	7,00

Liste des cooccurrents spécifiques de *il me semble que* dans l'empan gauche de 4 items (tableau de gauche) et dans l'empan droit de 10 items (tableau de droite) dans le corpus LM :

Occ	Fréq	CoFréq	Indice	DistMoy
SENT_?	26671	8	9,05	0,50
NAM_Saint-Pierre-des-Corps	13	2	8,03	2,00
SENT_.	668464	25	6,97	0,56
NOM_irresponsabilité	66	2	6,59	1,00
ADV_naturellement	353	2	5,13	2,00
NOM_cliché	401	2	5,02	1,00
ADJ_négatif	607	2	4,66	1,00
KON_or	977	2	4,25	0,00
ADV_au-delà	1782	2	3,73	3,00
NOM_forme	3350	2	3,18	3,00

Occ	Fréq	CoFréq	Indice	DistMoy
NOM_volantea	2	2	9,13	4,00
ADV_immensément	11	2	7,39	4,00
VER:cond_participer	28	2	6,55	8,00
NOM_bombardier	37	2	6,30	8,00
NOM_forteresse	71	2	5,73	3,00
ADV_activement	109	2	5,36	9,00
ADJ_aigu	125	2	5,24	8,00
VER:ppe_désigner	527	2	3,99	6,00
VER:pres_reprendre	623	2	3,84	6,00
VER:pres_devoir	11048	4	3,57	6,00
PRO:DEM_ce	159959	13	3,23	3,23
NOM_solidarité	1349	2	3,18	1,00
NOM_fait	6106	3	3,17	4,33
VER:infi_tenir	1390	2	3,15	7,00
ADJ_riche	1457	2	3,11	5,00
NOM_sentiment	1567	2	3,05	7,00
PRO:PER_je	26613	5	3,03	4,60

Je suis sûr·e que

Liste des cooccurrents spécifiques de *je suis sûr·e que* dans l'empan gauche de 4 items (tableau de gauche) et dans l'empan droit de 10 items (tableau de droite) dans le corpus PRES-FR :

Occ	Fréq	CoFréq	Indice	DistMoy
KON_et	55986	28	8,89	0,39
SENT_.	101711	39	8,88	0,41
VER:ppe_r_laisser	145	2	3,78	3,00

Occ	Fréq	CoFréq	Indice	DistMoy
VER:futu_arriver	78	3	5,63	2,33
VER:futu_approuver	10	2	5,35	2,00
NAM_Français	3564	8	4,64	2,00
VER:futu_être	3745	8	4,49	2,63
NOM_incident	40	2	4,11	4,50
PUN:cit_"	5498	9	4,09	4,11
VER:futu_savoir	47	2	3,97	5,00
VER:futu_faire	875	4	3,72	2,50
PRO:PER_vous	14325	14	3,60	2,57
VER:pres_comprendre	496	3	3,25	6,67
ADV_beaucoup	3239	6	3,18	4,17
VER:pres_préférer	132	2	3,08	4,50

Liste des cooccurrents spécifiques de *je suis sûr·e que* dans l'empan gauche de 4 items (tableau de gauche) et dans l'empan droit de 10 items (tableau de droite) dans le corpus LM :

Occ	Fréq	CoFréq	Indice	DistMoy
PUN:cit_«	157502	12	7,98	0,17
SENT_.	668464	20	6,12	1,20
ADV_aujourd'hui	8312	4	6,07	1,00
NOM_palestinien	496	2	5,08	1,00
NOM_porte	1607	2	4,06	1,50
PRP_contre	15251	3	3,52	3,00
KON_mais	46185	4	3,20	0,25

Occ	Fréq	CoFréq	Indice	DistMoy
PRO:PER_il	146572	17	7,71	3,82
VER:futu_trouver	281	3	7,50	1,67
PUN:cit_»	158082	13	4,47	5,38
PRO:PER_tu	977	2	3,70	0,00
PRO:PER_je	26613	5	3,58	2,80
NOM_peur	1619	2	3,26	2,50
NAM_Israël	1689	2	3,23	6,00
PRO:PER_me	8020	3	3,18	2,67
NOM_territoire	2021	2	3,07	3,00

Je suis convaincu·e que

Liste des cooccurrents spécifiques de *je suis convaincu·e que* dans l'empan gauche de 4 items (tableau de gauche) et dans l'empan droit de 10 items (tableau de droite) dans le corpus PRES-FR :

Occ	Fréq	CoFréq	Indice	DistMoy
SENT_.	101711	39	12,60	0,33
ADJ_religieux	100	2	4,35	1,50
NOM_reste	288	2	3,43	1,00
ADJ_culturel	340	2	3,29	3,00

Occ	Fréq	CoFréq	Indice	DistMoy
VER:futu_trouver	110	3	5,55	1,33
ADJ_religieux	100	2	3,56	8,50
VER:futu_rester	148	2	3,22	5,50
NOM_mal	175	2	3,08	3,50
NOM_partenaire	765	3	3,06	4,67

Liste des cooccurrents spécifiques de *je suis convaincu·e que* dans l'empan gauche de 4 items (tableau de gauche) et dans l'empan droit de 10 items (tableau de droite) dans le corpus LM :

Occ	Fréq	CoFréq	Indice	DistMoy
PUN:cit_«	157502	7	3,73	0,00
KON_parce	4304	2	3,31	1,00

Occ	Fréq	CoFréq	Indice	DistMoy
VER:pres_avancer	421	2	4,53	4,50
PRO:PER_vous	9937	3	3,05	6,33

Credo che

Liste des cooccurrents spécifiques de *credo che* dans l'empan gauche de 4 items (tableau de gauche) et dans l'empan droit de 10 items (tableau de droite) dans le corpus PRES-IT :

Occ	Fréq	CoFréq	Indice	DistMoy
SENT_.	80942	374	156,12	0,51
PRO:pers_io	1265	63	87,18	0,05
SENT_?	1609	29	27,92	1,17
ADV_però	760	12	11,33	0,50
CON_ma	7782	30	10,79	0,70
NOM_lo	30	4	7,88	0,25
PON_:	9750	23	4,83	0,22
CON_e	84212	108	4,82	0,95
ADV_francamente	57	3	4,78	0,67
NPR_R	61	3	4,69	1,33
NPR_Europa	6376	17	4,37	1,65
NOM_scommessa	14	2	4,20	1,00
NPR_Napolitano	1157	7	4,17	1,00
ADV_insomma	125	3	3,76	0,67
ADV_non	18039	31	3,67	0,52
ADV_no	144	3	3,58	1,33
ADV_ebbene	152	3	3,51	1,00
ADJ_altro	2071	8	3,36	1,38
NOM_presidente	7158	16	3,31	1,00
NOM_partita	43	2	3,21	0,50
ADJ_secondario	50	2	3,08	1,50

Occ	Fréq	CoFréq	Indice	DistMoy
VER:cpre_essere	2897	133	126,55	1,82
VER:cpre_potere	1271	55	51,45	2,65
CON_sia	1768	50	37,99	2,84
VER:cpre_dovere	407	30	35,36	2,13
PRO:demo_questo	19964	135	30,03	2,54
VER:cpre_avere	716	31	29,40	2,81
ADV_veramente	299	16	17,01	1,94
PRO:pers_si	7220	49	11,49	1,94
ADV_molto	2678	27	10,22	4,11
VER:infi_dire	1508	20	9,75	3,05
ADV_anche	11367	60	9,54	4,02
PRO:pers_noi	2602	25	9,10	2,72
ADJ_importante	2623	25	9,03	3,76
VER:cpre_andare	120	7	8,07	0,71
PRO:pers_ci	378	10	7,94	2,10
ADVdavvero	548	11	7,44	3,36
VER:pres_bisognare	762	12	6,92	1,25
ADV_non	18039	72	6,40	3,43
PRO:indef_qualcosa	322	8	6,27	1,88
PRO:indef_tutto	3805	25	5,99	2,84
CON_se	2893	21	5,80	3,05
VER:infi_essere	4455	27	5,76	3,63
PRO:indef_un	271	7	5,67	3,71
VER:futu_avere	400	8	5,57	3,63
VER:pper_essere stare	4875	28	5,51	3,93
PRO:demo_cio	1294	13	5,28	1,54
VER:cpre_stare	119	5	5,22	1,60
DET:indef_un	23505	83	5,22	4,20
VER:impe_po	233	6	4,93	3,67
PRO:pers_voi	1426	13	4,83	2,31
ADJ_giusto	691	9	4,70	2,56
ADV_seriamente	83	4	4,51	3,50
NOM_bisogno	1042	10	4,04	4,00
ADJ_nessuno	215	5	3,99	2,60
ADV_ragionevolmente	8	2	3,92	2,50
VER:pres_tracciare	8	2	3,92	0,50
VER:pper_soddisfare	48	3	3,83	2,33
NOM_sanitario	9	2	3,81	2,50
VER:cpre_collocare	9	2	3,81	2,50
VER:infi_ritenere	50	3	3,78	2,33
VER:cpre_costituire	54	3	3,68	4,00
ADV_ben	1178	10	3,62	3,20
NOM_accordo	573	7	3,61	4,14
VER:pres_potere	5712	26	3,55	2,35
VER:futu_passare	14	2	3,41	0,00
PRO:refl_ci	4593	22	3,39	2,55
PRE_in	33989	101	3,35	3,88
ADV_fortemente	301	5	3,32	2,80
ADV_altro	18	2	3,18	1,50
PRO:indef_nessuno	191	4	3,12	1,50
ADV_oggi	4516	21	3,11	3,76
NOM_parte	3294	17	3,10	4,18
NOM_c'è	1140	9	3,09	3,78
VER:infi_ricavare	20	2	3,09	3,00
PRO:indef_nessun	204	4	3,02	1,75

Liste des cooccurrents spécifiques de *credo che* dans l'empan gauche de 4 items (tableau ci-dessous) et dans l'empan droit de 10 items (tableau à la page suivante) dans le corpus REP :

Occ	Fréq	CoFréq	Indice	DistMoy
PRO:pers_io	9564	162	212,92	0,15
SENT_.	1064900	915	153,83	0,82
ADV_non	179623	351	152,25	0,20
CON_ma	94251	236	124,16	0,37
PON_"	516599	515	107,13	0,64
SENT_?	33142	110	70,26	1,07
PON_:	183087	204	49,34	1,10
PON_-	148959	148	30,83	1,14
CON_e	549784	316	21,35	0,74
PRO:demo_questo	82344	72	12,80	1,56
CON_perché	20923	32	11,83	0,19
ADV_personalmente	495	8	11,09	0,75
ADV_perché	6450	17	10,06	0,24
NOM_Non	2684	11	8,68	0,09
PON_«	12686	21	8,62	0,52
ADV_molto	19079	25	8,10	2,04
VER:pres_aggiungere	4032	12	7,87	1,75
ADJ_tanto	6452	13	6,53	2,00
NOM_“	23695	25	6,35	0,40
NOM_scudetto	2155	8	6,17	1,38
ADJ_competente	574	5	5,83	1,00
NOM_squadra	24351	24	5,62	1,96
NOM_partita	16591	19	5,49	1,53
ADV_bene	12342	16	5,38	1,06
ADV_quindi	9971	14	5,17	0,29
ADJ_tecnico	8923	13	5,02	1,69
ADV_francamente	123	3	5,02	0,00
PRO:poss_nostro	21368	21	4,99	2,19
ADJ_aggiunto	2368	7	4,84	1,86
ADV_però	16891	18	4,83	0,67
NPR_Milan	5687	10	4,68	1,70
ADV_sinceramente	170	3	4,60	1,33
NOM_fastidio	509	4	4,59	1,25
VER:cond_accettare	26	2	4,48	1,00
NOM_ragione	3076	7	4,12	1,14
NOM_giocatore	11000	13	4,10	2,08
ADV_discretamente	41	2	4,08	1,00
ADJ_importante	12669	14	4,06	1,79
NPR_Crespo	264	3	4,03	2,00
NOM_idea	5606	9	3,97	2,11
ADV_volentieri	286	3	3,93	0,67
NOM_Ghiraldini	63	2	3,70	1,50
NOM_problema	10732	12	3,61	2,17
ADJ_catastrofico	73	2	3,58	0,50
ADV_veramente	983	4	3,50	2,00
PON_...	8214	10	3,39	1,60
NOM_carisma	92	2	3,38	1,50
VER:pres_dimettere	95	2	3,35	0,50
ADV_poi	32440	23	3,31	0,61
VER:infi_migliorare	1968	5	3,31	1,60
VER:pper_aggiungere	2017	5	3,26	1,60
ADV_onestamente	111	2	3,21	0,50
VER:pres_piacere	2073	5	3,21	1,60
NOM_contrario	1196	4	3,18	0,25
VER:impf_vincere	118	2	3,16	1,00
NOM_mentalità	534	3	3,14	1,33
NOM_allenatore	5908	8	3,12	2,00
ADJ_facile	3290	6	3,11	1,83
ADJ_fiducioso	548	3	3,11	1,00
NOM_lo	551	3	3,10	1,00
ADJ_concentrato	130	2	3,08	1,50
NOM_Domizzi	132	2	3,06	1,00
ADV_no	4707	7	3,02	1,43

Occ	Fréq	CoFréq	Indice	DistMoy
VER:cpre avere	4397	192	255,46	2,30
VER:cpre dovere	1302	105	168,64	2,79
VER:cpre potere	5324	146	165,15	2,41
VER:cpre essere	15045	528	127,00	2,15
CON sia	7084	115	104,83	3,40
PRO:demo questo	82344	281	85,86	2,86
PRO:refl ci	37333	152	56,09	2,47
VER:cpre stare	831	26	31,68	1,58
ADJ giusto	5073	43	28,45	3,88
ADV molto	19079	76	27,97	4,32
NOM squadra	24351	85	27,31	3,19
VER:infi fare	34785	99	25,04	4,42
VER:futu essere	33988	97	24,64	2,35
ADV bene	12342	57	24,22	4,56
PRO:pers si	62674	138	23,97	2,70
NOM cosa	14815	60	22,64	4,40
VER:cpre volere	1537	24	22,24	2,33
PRO:pers ci	3535	32	22,24	1,47
VER:infi essere	31833	89	22,12	3,74
ADV non	179623	270	20,67	4,06
DET:def il	1546333	1568	20,52	3,36
VER:pper fare	24279	73	20,00	4,08
VER:cpre andare	963	19	19,64	3,05
PRO:pers noi	13427	53	19,63	3,96
DET:indef un	246346	330	17,72	3,91
VER:cimp meritare	22	7	16,46	1,00
PRO:indef tutto	19483	58	15,90	3,67
NOM partita	16591	53	15,82	4,55
VER:pper meritare	745	15	15,80	3,33
VER:futu riuscire	608	14	15,59	1,50
PRO:pers lo	24369	65	15,46	3,15
VER:futu fare	4461	27	14,66	2,52
VER:futu potere	6253	30	13,59	2,47
PRO:poss nostro	21368	56	13,14	4,02
ADJ importante	12669	41	12,62	4,80
NOM giocatore	11000	38	12,59	4,58
ADJ migliore	4995	25	11,85	4,00
PRO:poss mio	16220	44	10,98	4,18
VER:futu avere	4165	22	10,97	2,00
VER:infi giocare	5706	25	10,64	4,40
VER:cpre servire	171	7	9,86	2,29
PRO:demo quello	59301	97	9,79	4,89
NOM momento	14170	38	9,46	4,16
PRO:pers lui	12519	35	9,24	3,26
NOM problema	10732	32	9,17	3,72
VER:infi dare	8097	27	8,86	5,00
VER:cpre esistere	131	6	8,83	0,83
VER:cond essere	13923	36	8,60	2,47
ADJ buono	6769	24	8,46	4,38
PRO:indef un	596	21	8,18	3,48
PRO:indef qualcosa	4442	19	8,11	2,68
VER:cpre occorrere	29	4	8,02	3,00
ADV davvero	4099	18	7,89	3,50
ADV più	78862	112	7,82	4,38
VER:cpre meritare	37	4	7,58	4,25
VER:cpre sapere	570	8	7,53	3,63
VER:cimp essere	5383	20	7,50	2,40
NOM modo	10089	28	7,48	4,36
NOM società	10740	29	7,47	3,90
VER:pper sbagliare	1446	11	7,39	3,09
VER:cpre fare	1815	12	7,34	2,92
PRO:pers mi	20826	43	7,31	4,81
NOM calcio	7248	23	7,28	4,43
VER:impe po	5048	19	7,25	3,84
NPR Milan	5687	20	7,12	4,10
ADJ ottimo	2722	14	7,11	5,21
VER:futu rimanere	660	8	7,05	2,00
VER:geru fare	3700	16	7,00	3,56
VER:cimp potere	948	9	6,96	2,00
DET:indef una	168620	199	6,93	4,13
VER:pper equilibrare	471	7	6,84	5,14
NOM tipo	2938	14	6,71	4,21
NOM bisogno	3904	16	6,70	3,25
PRO:refl se	2957	14	6,68	0,21
PRO:indef nessuno	1715	11	6,65	0,82
PRO:indef qualcuno	3011	14	6,58	3,00
NOM fine	11968	29	6,51	4,52
ADV anche	91717	120	6,49	4,23
VER:futu vedere	2204	12	6,45	3,92
NOM errore	3647	15	6,34	3,67
VER:futu dovere	6405	20	6,32	2,85
VER:cimp avere	1872	11	6,28	1,82
VER:impe demeritare	2	2	6,20	3,50
VER:impe avere	202	5	6,12	2,80

NOM ortodossia	22	3	6,11	4,00
VER:futu lottare	85	4	6,11	0,25
VER:infi migliorare	1968	11	6,07	5,00
ADV oggi	22491	42	6,05	2,57
CON se	36768	59	6,02	4,73
PRE come	27516	48	5,98	4,60
VER:infi tornare	3398	14	5,97	4,21
VER:infi vincere	5017	17	5,94	4,35
ADJ qualche	10085	25	5,89	3,36
ADJ nessuno	3976	15	5,88	3,13
NOM lbra	227	5	5,87	1,80
ADV ci	231	5	5,83	3,60
ADJ buon	3020	13	5,80	3,31
ADJ certo	4088	15	5,74	3,13
VER:impe arrivera	3	2	5,72	0,50
VER:cond bisognare	244	5	5,72	2,00
ADV perché	6450	19	5,68	5,37
ADV ancora	29108	49	5,67	4,24
VER:cpre dipendere	33	3	5,56	0,33
ADJ necessario	4818	16	5,53	2,56
VER:futu vincere	272	5	5,49	1,20
PRO:pers me	6676	19	5,47	5,63
ADV ce	2750	12	5,46	2,00
ADV così	21676	39	5,30	4,46
DET:indef un'	29220	48	5,30	3,88
NOM gente	3370	13	5,29	4,00
VER:pres credere	4477	15	5,27	6,13
NOM rosa	1153	8	5,25	2,88
VER:cpre compensare	5	2	5,20	3,50
ADV meglio	6351	18	5,19	4,06
NOM possibilità	5183	16	5,14	4,38
VER:futu dare	1201	8	5,13	2,38
VER:infi dire	8428	21	5,09	4,90
ADJ allucinante	49	3	5,04	0,67
VER:futu abituare	6	2	5,02	0,50
VER:infi parlare	4137	14	5,01	3,79
ADV tutto	7880	20	5,00	2,10
VER:cond potere	10677	24	4,99	2,58
VER:cpre rimanere	169	4	4,92	3,50
VER:pres dovere	19493	35	4,81	4,23
VER:pper condizionare	377	5	4,80	3,40
VER:infi lottare	648	6	4,77	4,17
VER:cpre piacere	60	3	4,77	0,67
VER:pres potere	26362	43	4,75	4,07
ADJ grande	30731	48	4,75	4,85
VER:infi ripetere	680	6	4,66	3,67
VER:futu lasciare	411	5	4,63	2,20
PRO:demo ciò	3373	12	4,59	3,00
CON visto	3421	12	4,53	5,17
VER:cond avere	10782	23	4,46	2,13
NOM risultato	10907	23	4,39	4,39
VER:futu sapere	462	5	4,39	2,60
VER:pper dimostrare	2006	9	4,34	5,11
VER:futu accadere	238	4	4,34	3,75
VER:futu andare	2508	10	4,33	3,60
ADV quando	3651	12	4,27	1,42
NOM Mourinho	1179	7	4,25	3,43
VER:futu superare	90	3	4,25	0,67
CON perché	20923	35	4,21	5,89
VER:pper vivere	1197	7	4,21	5,71
VER:cond dovere	6264	16	4,16	3,88
ADJ accettabile	271	4	4,12	5,25
VER:futu giocare	1251	7	4,09	2,29
VER:pper legare	2703	10	4,07	5,50
ADJ difficile	7137	17	4,02	4,94
ADJ corresponsabile	18	2	4,01	1,50
VER:futu arrivare	1739	8	4,00	3,13
VER:infi stare	1751	8	3,98	3,88
PRO:pers la	5843	15	3,97	1,33
VER:infi avere	21695	35	3,91	3,77
NOM questione	3991	12	3,91	3,33
ADJ opportuno	591	5	3,89	3,20
PRO:indef ognuno	948	6	3,87	2,50
VER:futu avvenire	316	4	3,87	0,50
VER:infi aiutare	1361	7	3,86	4,71
ADJ fondamentale	2872	10	3,86	4,40
NOM Brawn	22	2	3,84	4,00
NOM elemento	2348	9	3,83	3,89
VER:futu attaccare	23	2	3,80	3,50
NOM decisione	5397	14	3,79	3,86
ADJ sufficiente	981	6	3,79	3,17
NOM Mou	129	3	3,78	2,67
NPR Champions	3528	11	3,77	5,18
ADV sicuramente	2399	9	3,76	4,33
VER:futu reggere	24	2	3,76	0,50

NPR Juve	4179	12	3,73	4,17
VER:infi perdere	2450	9	3,70	4,44
VER:pres stare	19838	32	3,64	5,03
VER:futu finire	370	4	3,61	0,50
VER:pper influenzare	149	3	3,60	1,67
VER:cpre riflettere	29	2	3,59	2,50
VER:infi aspettare	1078	6	3,57	2,00
NOM importante	156	3	3,54	2,00
ADJ deleterio	32	2	3,51	4,00
VER:futu prevalere	32	2	3,51	2,00
VER:cpre vincere	162	3	3,49	3,67
VER:cpre valere	34	2	3,45	0,00
VER:futu perdere	167	3	3,45	1,33
VER:geru succedere	168	3	3,45	3,33
ADJ normale	1606	7	3,43	2,71
NPR Inter	6574	15	3,43	2,67
PRO:pers tu	412	4	3,43	1,50
NOM difficoltà	4516	12	3,43	3,50
NOM salvezza	1150	6	3,42	4,33
VER:infi pensare	3275	10	3,41	3,50
VER:infi puntare	756	5	3,40	3,80
NOM senso	3910	11	3,39	4,45
VER:cpre dire	180	3	3,36	4,00
VER:infi condizionare	180	3	3,36	3,00
NOM dov'è	38	2	3,36	2,50
NPR Moratti	1676	7	3,32	2,14
VER:geru giocare	786	5	3,32	2,60
NOM veto	187	3	3,31	2,00
ADJ simile	1694	7	3,30	4,71
ADV forse	6103	14	3,26	4,93
ADJ competitivo	810	5	3,26	5,20
VER:cpre rischiare	43	2	3,25	0,00
NOM qualità	4738	12	3,25	4,17
VER:futu tornare	1253	6	3,23	0,67
VER:futu recuperare	45	2	3,21	0,50
NOM riflessione	848	5	3,17	3,60
ADJ valgo	47	2	3,17	1,50
ADJ accetto	213	3	3,15	5,67
VER:infi mettere	5564	13	3,15	3,00
VER:infi trovare	4881	12	3,14	4,92
VER:cpre arrivare	498	4	3,12	1,75
ADJ aggiunto	2368	8	3,12	5,75
NOM futuro	5614	13	3,11	5,08
NOM credo	509	4	3,09	6,00
NOM rispetto	7883	16	3,08	5,44
NOM valore	4957	12	3,08	4,50
ADV ne	12784	22	3,04	3,09
NOM dovere	910	5	3,04	2,40
ADJ tanto	6452	14	3,04	5,79
ADJ antipatico	55	2	3,04	2,50
VER:infi vivere	1885	7	3,03	3,86
ADV stasera	2446	8	3,03	3,75
ADJ maturo	236	3	3,02	2,00
NOM mentalità	534	4	3,01	5,50
NOM mister	922	5	3,01	0,80
VER:futu stare	239	3	3,00	2,67

Penso che

Liste des cooccurrents spécifiques de *penso che* dans l'empan gauche de 4 items (tableau de gauche) et dans l'empan droit de 10 items (tableau de droite) dans le corpus PRES-IT :

Occ	Fréq	CoFréq	Indice	DistMoy
SENT_.	80942	145	58,98	0,49
PRO:pers_io	1265	24	33,05	0,08
SENT_?	1609	18	20,90	1,11
NOM_io	30	6	15,06	0,00
NOM_presidente	7158	14	6,58	0,79
CON_perché	2666	7	4,37	0,14
NOM_minuto	28	2	4,37	1,00
CON_ma	7782	10	3,40	0,20
NPR_Napolitano	1157	4	3,15	1,25

Occ	Fréq	CoFréq	Indice	DistMoy
VER:cpre_essere	2897	37	30,04	1,81
VER:cpre_dovere	407	18	24,51	2,83
VER:cpre_potere	1271	21	19,64	2,81
PRO:demo_questo	19964	53	11,96	2,36
ADV_molto	2678	15	7,76	4,47
VER:infi_essere	4455	17	6,34	3,82
CON_sia	1768	11	6,31	2,55
ADJ_buono	406	6	5,80	3,00
VER:cpre_avere	716	7	5,48	2,14
VER:cpre_andare	120	4	5,42	3,00
VER:pres_bisognare	762	7	5,31	2,29
ADV_forse	600	6	4,84	4,17
VER:cpre_costruire	7	2	4,83	0,50
NOM_gusto	72	3	4,47	4,67
VER:cpre_servire	13	2	4,26	1,50
VER:cimp_avere	101	3	4,03	2,67
VER:cpre_stare	119	3	3,82	1,67
VER:infi_riaprire	22	2	3,79	1,00
PRO:indef_qualcosa	322	4	3,75	4,75
VER:cpre_esistere	24	2	3,71	0,50
CON_se	2893	10	3,65	3,40
ADV_anche	11367	22	3,42	3,64
VER:futu_potere	730	5	3,36	3,20
PRO:pers_noi	2602	9	3,35	1,78
NOM_contributo	1212	6	3,19	5,17
NOM_filone	44	2	3,18	2,50
ADJ_nessuno	215	3	3,07	0,33

Liste des cooccurrents spécifiques de *penso che* dans l'empan gauche de 4 items (tableau ci-dessous) et dans l'empan droit de 10 items (tableau à la page suivante) dans le corpus REP :

Occ	Fréq	CoFréq	Indice	DistMoy
PRO:pers_io	9564	88	116,54	0,15
SENT_.	1064900	481	78,65	0,75
PON_"	516599	296	68,51	0,66
CON_ma	94251	112	54,30	0,56
SENT_?	33142	63	42,33	1,16
PON_:	183087	121	33,64	1,07
ADV_non	179623	98	21,49	0,41
PON_-	148959	71	12,99	1,18
CON_e	549784	162	10,00	0,59
NOM_io	551	6	9,05	0,00
PRO:demo_questo	82344	42	8,84	1,74
ADV_cosi	21676	19	7,78	1,26
ADV_si	1894	7	7,25	1,29
NOM_“	23695	18	6,51	0,28
ADV_bene	12342	13	6,41	1,15
CON_se	36768	22	6,06	1,00
ADV_onestamente	111	3	5,96	0,33
ADV_quindi	9971	11	5,72	0,36
ADV_perché	6450	9	5,60	0,00
PON_«	12686	11	4,75	0,45
ABR_Cuperlo	46	2	4,52	0,50
NOM_squadra	24351	15	4,49	1,93
VER:infi_vincere	5017	7	4,48	2,29
PRO:poss_mio	16220	12	4,45	2,00
NOM_calcio	7248	8	4,33	1,63
CON_perché	20923	13	4,01	0,38
VER:pper_dire	15357	11	4,01	2,09
VER:pper_deludere	514	3	3,97	1,00
PRO:refl_se	2957	5	3,74	0,40
NOM_partita	16591	11	3,73	1,09
NOM_arbitraggio	154	2	3,47	0,50
NOM_guardo	177	2	3,35	1,50
VER:infi_stupire	182	2	3,32	0,50
VER:pres_piacere	2073	4	3,30	1,25
NOM_problema	10732	8	3,19	1,88
VER:geru_accadere	234	2	3,11	1,00
ADV_però	16891	10	3,07	0,20
NOM_Cota	253	2	3,04	1,00

Occ	Fréq	CoFréq	Indice	DistMoy
VER:cpre_essere	15045	245	287,19	1,67
VER:cpre_avere	4397	90	114,84	2,63
VER:cpre_dovere	1302	47	72,00	2,62
VER:cpre_potere	5324	66	70,37	2,29
CON_sia	7084	55	48,06	3,56
PRO:demo_questo	82344	117	26,93	3,33
PRO:refl_ci	37333	70	22,76	2,81
VER:infi_essere	31833	60	19,75	4,25
VER:infi_fare	34785	61	18,58	4,38
NOM_cosa	14815	40	18,56	4,78
VER:cpre_stare	831	13	15,82	1,85
VER:pper_fare	24279	46	15,43	3,52
PRO:pers_ci	3535	19	14,32	0,74
VER:cpre_andare	963	12	13,50	3,75
ADV_bene	12342	30	12,99	4,80
ADV_più	78862	82	11,86	4,54
VER:futu_fare	4461	18	11,56	2,06
VER:futu_essere	33988	48	11,49	2,44
PRO:poss_mio	16220	32	11,44	4,22
ADJ_importante	12669	28	11,22	3,79
NOM_squadra	24351	39	11,04	4,13
NOM_giocatore	11000	25	10,35	3,84
DET:def_il	1546333	829	10,32	3,55
VER:cond_dovere	6264	19	10,07	2,05
PRO:pers_noi	13427	27	9,95	4,96
VER:cpre_volere	1537	11	9,88	2,00
ADJ_giusto	5073	17	9,73	3,35
VER:futu_giocare	1251	10	9,50	1,40
ADV_non	179623	137	9,49	4,23
DET:indef_un	246346	174	9,33	4,40
VER:infi_vincere	5017	16	8,90	4,19
PRO:pers_lo	24369	35	8,79	3,89
VER:cpre_meritare	37	4	8,66	1,25
VER:cond_essere	13923	25	8,31	1,96
ADJ_buon	3020	12	7,87	4,67
NOM_ragione	3076	12	7,79	4,08
ADV_anche	91717	78	7,54	4,23
NOM_partita	16591	26	7,44	4,58
VER:cpre_mancare	74	4	7,43	0,75
NOM_gente	3370	12	7,36	3,58
PRO:pers_me	6676	16	7,20	4,69
ADJ_buono	6769	16	7,12	4,38
PRO:poss_nostro	21368	29	6,91	4,45
PRO:indef_tanto	6162	15	6,88	2,67
VER:infi_vedere	5381	14	6,81	3,93
PRO:pers_si	62674	57	6,60	2,63
VER:cpre_occorrere	29	3	6,55	3,00
VER:cpre_dipendere	33	3	6,37	0,33
VER:pres_pensare	5962	14	6,29	6,36
VER:pres_bisognare	4281	12	6,27	1,75
NOM_scudetto	2155	9	6,25	5,22
VER:pres_dovere	19493	26	6,13	4,19
PRO:refl_se	2957	10	6,04	1,60
VER:geru_fare	3700	11	6,04	3,64
VER:infi_giocare	5706	13	5,74	3,38
PRO:indef_tutto	19483	25	5,64	3,80
ADJ_migliore	4995	12	5,58	3,00
VER:futu_avere	4165	11	5,55	3,64
DET:indef_una	168620	114	5,53	4,50
PRO:pers_lui	12519	19	5,44	3,05
NOM_possibilità	5183	12	5,42	5,42
VER:futu_potere	6253	13	5,32	3,38
NOM_fine	11968	18	5,14	3,33
ADV_tutto	7880	14	4,92	2,21
VER:futu_arrivare	1739	7	4,89	1,14

VER:futu_dare	1201	6	4,79	2,00
NOM_DRS	14	2	4,78	1,50
VER:futu_andare	2508	8	4,78	1,50
NPR_Juve	4179	10	4,74	3,80
VER:infi_cambiare	2538	8	4,74	4,75
VER:pper_meritare	745	5	4,68	1,60
VER:pres_stare	19838	23	4,56	3,35
ADV_ce	2750	8	4,50	1,00
PRO:demo_quello	59301	48	4,41	4,31
VER:cimp_meritare	22	2	4,38	1,50
NOM_livello	6725	12	4,32	5,58
VER:pres_potere	26362	27	4,32	2,41
NOM_esperienza	3792	9	4,30	4,11
VER:pper_capire	1485	6	4,28	3,33
PRO:pers_mi	20826	23	4,25	4,17
VER:cpre_permettere	168	3	4,24	1,33
VER:cpre_servire	171	3	4,22	1,67
VER:futu_influire	28	2	4,16	0,00
PRO:pers_loro	7013	12	4,15	3,08
ADJ_nessuno	3976	9	4,15	5,00
VER:pper_migliorare	498	4	4,15	1,50
ADVdavvero	4099	9	4,05	2,78
NOM_problema	10732	15	4,05	3,60
CON_se	36768	33	4,03	4,33
VER:cond_avere	10782	15	4,02	2,00
PRO:pers_io	9564	14	4,02	6,21
NOM_tifoso	5194	10	3,97	2,50
VER:futu_mettere	572	4	3,92	2,00
ADV_meglio	6351	11	3,90	3,27
VER:geru_valorizzare	41	2	3,83	2,00
VER:pper_soddisfare	1827	6	3,79	3,67
VER:pres_dire	15328	18	3,78	5,78
ADV_qui	8998	13	3,73	2,15
NOM_momento	14170	17	3,71	3,47
VER:infi_mandare	679	4	3,63	3,50
ADJ_fondamentale	2872	7	3,54	2,71
ADV_stare	718	4	3,54	4,75
ADV_molto	19079	20	3,50	4,15
VER:futu_uscire	304	3	3,48	2,00
VER:futu_volere	749	4	3,47	3,75
ADJ_doveroso	315	3	3,44	0,67
NOM_ascoltatore	67	2	3,40	2,50
PRO:indef_molto	6195	10	3,37	1,20
VER:cpre_parlare	335	3	3,36	0,33
ADJ_competitivo	810	4	3,34	2,50
NOM_prestazione	3129	7	3,32	4,86
VER:cpre_vedere	347	3	3,32	0,67
VER:pper_avere	7483	11	3,31	4,55
ADV_troppo	5253	9	3,27	3,56
NOM_modo	10089	13	3,26	4,69
ADV_sempre	22977	22	3,26	3,86
VER:futu_mandare	82	2	3,23	3,00
VER:futu_lottare	85	2	3,20	0,00
ADJ_molto	4323	8	3,18	4,38
NOM_mister	922	4	3,14	0,50
VER:pper_reagire	405	3	3,12	1,33
NOM_stia	93	2	3,12	3,00
VER:futu_rivedere	97	2	3,08	2,50
VER:infi_soffrire	420	3	3,08	2,67
VER:pres_avere	255987	144	3,07	4,74
VER:infi_influenzare	100	2	3,06	2,50
PRO:pers_gli	6822	10	3,05	4,80
VER:infi_dare	8097	11	3,03	4,82
ADJ_soddisfatto	104	2	3,02	2,00
VER:futu_venire	4594	8	3,01	3,63
VER:infi_criticare	106	2	3,01	0,50

Ritengo che

Liste des cooccurrents spécifiques de *ritengo che* dans l'empan gauche de 4 items (tableau de gauche) et dans l'empan droit de 10 items (tableau de droite) dans le corpus PRES-IT :

Occ	Fréq	CoFréq	Indice	DistMoy
PRO:pers_io	1265	21	35,43	0,10
SENT_.	80942	60	23,84	0,60
CON_perché	2666	7	6,72	0,00
PON_)	1202	5	5,92	1,00

Occ	Fréq	CoFréq	Indice	DistMoy
VER:cpre_essere	2897	26	26,52	1,15
VER:cpre_dovere	407	9	13,10	2,78
VER:cpre_potere	1271	8	7,45	1,75
PRO:pers_si	7220	13	5,37	2,08
ADV_altro	18	2	4,69	2,00
PRO:demo_questo	19964	21	4,59	0,95
PRE_senz'	25	2	4,40	1,50
VER:cpre_esprimere	51	2	3,77	3,00
ADV_ancora	3539	7	3,40	4,43
CON_sia	1768	5	3,25	2,80

Liste des cooccurrents spécifiques de *ritengo che* dans l'empan gauche de 4 items (tableau de gauche) et dans l'empan droit de 10 items (tableau de droite) dans le corpus REP :

Occ	Fréq	CoFréq	Indice	DistMoy
PRO:pers_io	9564	14	19,59	0,29
PON_"	516599	49	13,68	0,53
SENT_.	1064900	67	10,32	0,90
PON_-	148959	20	8,33	1,25
ADV_pertanto	550	3	6,33	0,00
PON_:	183087	19	6,25	0,89
CON_perché	20923	6	4,66	0,00
ADV_altrove	501	2	4,08	1,00
CON_poiché	703	2	3,79	0,00
PRO:demo_questo	82344	9	3,44	1,33
ADJ_tranquillo	1315	2	3,25	0,50
NOM_perché	1518	2	3,12	0,00
NOM_riduzione	1693	2	3,03	0,50
CON_ma	94251	9	3,03	0,33
NOM_fenomeno	1717	2	3,02	0,50

Occ	Fréq	CoFréq	Indice	DistMoy
VER:cpre_essere	15045	53	71,04	1,98
VER:cpre_dovere	1302	22	44,86	2,41
VER:cpre_potere	5324	12	14,39	3,67
VER:cpre_avere	4397	11	13,74	2,82
CON_sia	7084	12	12,95	3,08
ADV_non	179623	38	9,22	3,03
VER:infi_essere	31833	16	9,16	4,00
NOM_tracotanza	6	2	7,20	0,50
VER:cpre_andare	963	4	6,23	4,00
VER:cpre_sussistere	18	2	6,19	0,50
VER:cpre_occorrere	29	2	5,77	0,00
ADJ_opportuno	591	3	5,05	1,33
PRO:demo_questo	82344	18	4,92	1,78
VER:cpre_servire	171	2	4,22	0,00
PRO:pers_si	62674	14	4,07	1,64
PRO:pers_ci	3535	4	4,02	0,50
DET:def_il	1546333	137	3,76	3,26
NOM_presupposto	298	2	3,74	2,00
DET:indef_un	246346	32	3,62	3,66
NOM_opportunità	1904	3	3,55	4,00
ADJ_giusto	5073	4	3,43	1,75
DET:indef_una	168620	24	3,40	3,38
VER:cpre_trattare	536	2	3,23	0,50
NOM_gestore	587	2	3,15	2,00
ADJ_eccessivo	605	2	3,13	3,00
NOM_delega	650	2	3,07	3,00

Sono convinto/a che

Liste des cooccurrents spécifiques de *sono convinto/a che* dans l'empan gauche de 4 items (tableau de gauche) et dans l'empan droit de 10 items (tableau de droite) dans le corpus PRES-IT :

Occ	Fréq	CoFréq	Indice	DistMoy
SENT_.	80942	119	53,80	0,37
PRO:pers_io	1265	7	7,36	0,14
SENT_?	1609	5	4,22	0,60
CON_ma	7782	9	3,75	0,44
CON_perché	2666	5	3,22	0,00

Occ	Fréq	CoFréq	Indice	DistMoy
VER:cpre_essere	2897	15	9,06	3,53
PRO:demo_questo	19964	39	9,02	2,64
VER:futu_essere	1455	11	8,48	3,73
DET:def_il	150004	152	8,40	3,42
VER:futu_potere	730	7	6,30	3,14
VER:cpre_andare	120	4	5,94	2,50
CON_sia	1768	9	5,64	1,44
ADJ_grande	6496	16	5,30	4,63
VER:futu_arrivare	9	2	4,86	1,00
VER:futu_fare	250	4	4,68	5,75
ADV_insieme	2465	9	4,51	5,44
VER:futu_riuscire	97	3	4,47	2,00
VER:futu_avvenire	25	2	3,94	3,50
VER:futu_aver	400	4	3,90	6,75
VER:cpre_potere	1271	6	3,78	1,83
VER:futu_continuare	186	3	3,64	0,00
NOM_magistrato	479	4	3,60	4,25
NOM_serenità	204	3	3,52	9,00
PRO:pers_voi	1426	6	3,51	3,83
ADV_anche	11367	18	3,43	2,61
ADV_veramente	299	3	3,04	3,67

Liste des cooccurrents spécifiques de *sono convinto/a che* dans l'empan gauche de 4 items (tableau de gauche) et dans l'empan droit de 10 items (tableau de droite) dans le corpus REP :

Occ	Fréq	CoFréq	Indice	DistMoy
CON_ma	94251	49	28,26	0,29
PRO:pers_io	9564	23	28,11	0,17
SENT_.	1064900	168	27,39	0,61
CON_e	549784	98	19,43	0,57
PON_:	183087	34	7,59	0,71
PON_"	516599	59	5,20	0,44
PON_-	148959	25	5,03	0,76
ADV_anzi	2269	4	4,87	0,50
ADJ_grande	30731	10	4,57	2,00
NPR_Uruguay	290	2	3,81	0,50
ADJ_forte	8108	5	3,79	1,40
NOM_meraviglia	338	2	3,68	1,00
VER:pres_conoscere	1979	3	3,57	1,00
CON_perché	20923	7	3,44	0,00
ADJ_disponibile	2364	3	3,35	1,33
NOM_Delneri	504	2	3,34	0,50
NOM_allenatore	5908	4	3,28	1,25
NOM_Marotta	565	2	3,24	1,50
NOM_ambizione	573	2	3,22	1,00
PRO:indef_tanto	6162	4	3,21	2,00
NOM_risultato	10907	5	3,20	1,00
NOM_serenità	640	2	3,13	0,50
NPR_Europa	6853	4	3,04	1,00

Occ	Fréq	CoFréq	Indice	DistMoy
VER:futu_fare	4461	27	32,90	2,15
VER:cpre_essere	15045	36	29,63	1,67
PRO:demo_questo	82344	61	22,15	3,62
VER:cpre_potere	5324	21	21,98	2,38
VER:infi_fare	34785	39	20,35	3,87
VER:futu_essere	33988	38	19,81	2,42
PRO:refl_ci	37333	35	16,02	2,51
VER:futu_dare	1201	11	15,81	2,27
VER:cpre_dovere	1302	11	15,43	2,45
VER:cpre_avere	4397	13	12,33	1,62
VER:futu_potere	6253	14	11,63	1,21
VER:futu_riuscire	608	7	10,98	1,86
VER:pres_potere	26362	24	10,98	1,83
VER:futu_uscire	304	6	10,90	1,17
ADV_bene	12342	17	10,67	3,35
VER:futu_sapere	462	6	9,81	3,17
CON_sia	7084	12	8,72	4,08
VER:cpre_andare	963	6	7,92	2,50
NOM_squadra	24351	19	7,80	3,58
NOM_momento	14170	14	7,13	5,64
CON_se	36768	22	6,91	3,14
VER:futu_trovare	823	5	6,64	2,00
NOM_appaluso	5	2	6,63	3,50
ADV_ce	2750	7	6,51	1,00
PRO:poss_mio	16220	14	6,44	3,93
PRO:pers_la	5843	9	6,36	3,56
NOM_partita	16591	14	6,32	5,36
ADJ_buon	3020	7	6,24	3,71
VER:infi_dare	8097	10	6,13	3,50
DET:def_il	1546333	306	5,45	3,31
ADV_non	179623	54	5,37	5,00
VER:infi_adattare	203	3	5,32	0,67
ADJ_forte	8108	9	5,21	4,89
DET:indef_una	168620	51	5,18	3,39
VER:futu_vincere	272	3	4,94	0,00
ADV_molto	19079	13	4,93	4,15
ADV_benissimo	962	4	4,77	3,50
NOM_pallavolo	322	3	4,72	2,33
VER:futu_scendere	336	3	4,67	0,67
ADV_più	78862	29	4,64	4,59
ADJ_grande	30731	16	4,49	3,69
ADJ_importante	12669	10	4,46	5,90
PRO:poss_nostro	21368	13	4,43	4,38
ADV_così	21676	13	4,37	3,15
VER:futu_tornare	1253	4	4,32	1,25
VER:infi_cambiare	2538	5	4,28	4,40
PRO:pers_noi	13427	10	4,25	3,30
NOM_milanese	468	3	4,24	0,67
ADV_ancora	29108	15	4,20	3,27
NOM_massimo	2639	5	4,20	3,40
VER:cimp_dovere	1420	4	4,11	2,75
NOM_grado	4668	6	4,01	5,00
NOM_soddisfazione	1592	4	3,92	4,50
NOM_cosa	14815	10	3,91	3,60
ADJ_giusto	5073	6	3,81	5,17
NOM_mano	7309	7	3,79	4,57
VER:infi_essere	31833	15	3,78	5,33
VER:futu_accogliere	121	2	3,78	1,50
VER:cpre_fare	1815	4	3,71	2,00
NOM_soluzione	3391	5	3,70	4,00
VER:futu_sentire	140	2	3,65	2,50
NOM_opportunità	1904	4	3,63	3,75
ADJ_competitivo	810	3	3,54	4,33
VER:infi_giocare	5706	6	3,54	2,00
NOM_Vinovo	159	2	3,54	2,50
VER:cpre_stare	831	3	3,51	2,67
NOM_campionato	8175	7	3,50	5,00
ADV_presto	2152	4	3,43	3,00
NOM_lega	3983	5	3,38	3,60
NOM_situazione	8640	7	3,36	4,14
VER:pper_unire	937	3	3,36	1,33
ADV_domani	6196	6	3,35	1,17
PON_"	516599	110	3,34	5,88
ADJ_meglio	227	2	3,23	3,00
ADJ_rigorouso	239	2	3,19	1,50
VER:futu_andare	2508	4	3,18	2,00
VER:futu_continuare	1093	3	3,17	0,00
VER:geru_cambiare	265	2	3,10	1,50
PRO:pers_lui	12519	8	3,09	4,13
NOM_Salvini	274	2	3,07	2,50

Mi pare che

Liste des cooccurrents spécifiques de *mi pare che* dans l'empan gauche de 4 items (tableau de gauche) et dans l'empan droit de 10 items (tableau de droite) dans le corpus PRES-IT :

Occ	Fréq	CoFréq	Indice	DistMoy
SENT_.	80942	41	14,46	0,59
SENT_?	1609	4	4,38	2,50
ADV_chiaramente	114	2	4,07	2,50
NOM_presidente	7158	6	3,68	1,33
CON_perché	2666	4	3,55	0,50
NPR_Napolitano	1157	3	3,45	0,33
VER:pper_riconoscere	310	2	3,21	3,00

Occ	Fréq	CoFréq	Indice	DistMoy
VER:cpre_essere	2897	33	40,25	1,88
VER:cpre_avere	716	7	8,60	1,71
PRO:demo_questo	19964	23	7,41	2,65
CON_sia	1768	7	5,96	3,00
ADV_molto	2678	8	5,82	4,75
PRO:pers_ci	378	4	5,28	1,00
NOM_senso	1809	6	4,76	4,17
NOM_punto	1371	5	4,25	3,20
VER:cpre_porre	46	2	4,07	6,00
VER:pper_essere stare	4875	8	3,98	2,25
VER:cpre_trattare	56	2	3,90	1,00
PRO:pers_egli	335	3	3,86	4,00
VER:cpre_dovere	407	3	3,61	1,33
VER:pper_mettere	456	3	3,47	5,00
NOM_c'è	1140	4	3,43	4,50
ADV_già	2194	5	3,31	3,00
ADV_non	18039	14	3,02	4,43

Liste des cooccurrents spécifiques de *mi pare che* dans l'empan gauche de 4 items (tableau de gauche) et dans l'empan droit de 10 items (tableau de droite) dans le corpus REP :

Occ	Fréq	CoFréq	Indice	DistMoy
ADV_non	179623	26	14,59	0,31
CON_ma	94251	14	8,29	0,57
SENT_.	1064900	49	8,00	0,82
SENT_?	33142	7	5,42	0,71
PON_"	516599	23	3,84	0,91

Occ	Fréq	CoFréq	Indice	DistMoy
VER:cpre_essere	15045	36	47,67	2,58
VER:cpre_avere	4397	22	36,39	2,41
VER:cpre_stare	831	7	13,62	2,00
NOM_stia	93	4	10,87	1,50
CON_sia	7084	8	8,60	3,38
ADV_bene	12342	8	6,76	3,13
PRO:pers_ci	3535	4	4,57	1,75
VER:pper_venire	1597	3	4,19	3,00
PRO:refl_ci	37333	9	4,08	1,67
VER:cimp_essere	5383	4	3,87	2,75
PRO:demo_questo	82344	13	3,75	1,08
VER:pper_meritare	745	2	3,23	0,50
VER:geru_fare	3700	3	3,13	2,33

Sono sicuro/a che

Liste des cooccurrents spécifiques de *sono sicuro/a che* dans l'empan gauche de 4 items (tableau de gauche) et dans l'empan droit de 10 items (tableau de droite) dans le corpus PRES-IT :

Occ	Fréq	CoFréq	Indice	DistMoy
SENT_.	80942	39	16,57	0,10
PRO:pers_io	1265	3	3,64	0,00
CON_e	84212	18	3,10	0,83

Occ	Fréq	CoFréq	Indice	DistMoy
VER:futu_sapere	215	6	10,78	3,00
VER:futu_fare	250	4	6,40	2,00
PRO:poss_vostro	3500	8	5,72	4,13
VER:futu_dare	103	3	5,69	2,67
DET:def_il	150004	63	5,55	3,29
PRO:demo_questo	19964	17	5,29	3,65
NOM_parte	3294	7	4,88	5,00
VER:futu_rappresentare	26	2	4,78	2,50
PRO:pers_voi	1426	5	4,65	0,60
VER:futu_tenere	33	2	4,57	3,00
VER:futu_potere	730	4	4,56	5,50
VER:futu_costituire	36	2	4,49	4,00
VER:futu_essere	1455	4	3,42	1,50
VER:futu_venire	141	2	3,30	4,50
ADJ_duraturo	177	2	3,11	8,50

Liste des cooccurrents spécifiques de *sono sicuro/a che* dans l'empan gauche de 4 items (tableau de gauche) et dans l'empan droit de 10 items (tableau de droite) dans le corpus REP :

Occ	Fréq	CoFréq	Indice	DistMoy
CON_ma	94251	44	28,33	0,09
CON_e	549784	79	17,21	0,34
SENT_.	1064900	108	12,98	0,40
ADV_bene	12342	9	7,97	1,22
PRO:pers_io	9564	8	7,60	0,38
PON_:	183087	27	6,53	0,81
PRO:demo_questo	82344	15	4,94	1,67
NOM_book	108	2	4,91	1,50
PRO:poss_mio	16220	7	4,85	2,00
NOM_perché	1518	3	4,26	0,33
NOM_momento	14170	6	4,19	2,17
VER:infi_restare	1615	3	4,18	2,00
ADV_no	4707	4	4,10	1,50
ADJ_automobilistico	455	2	3,66	0,50
NOM_tranquillità	493	2	3,59	0,50
VER:pper_vedere	6703	4	3,52	1,25
NPR_Olimpiadi	572	2	3,46	0,50
VER:futu_riuscire	608	2	3,41	0,00
ADJ_fotografico	749	2	3,23	1,00
NOM_ritorno	3669	3	3,14	0,33
NOM_industria	886	2	3,08	1,00
ADJ_sereno	942	2	3,03	0,50

Occ	Fréq	CoFréq	Indice	DistMoy
VER:futu_fare	4461	39	57,8299	1,7949
VER:futu_dare	1201	17	29,2146	2
VER:futu_essere	33988	43	28,919	3,4651
ADV_bene	12342	23	19,4457	3,5217
VER:futu_sapere	462	9	17,0505	1,6667
VER:futu_trovare	823	10	16,8208	0,9
VER:futu_riuscire	608	9	15,9755	1,7778
PRO:refl_ci	37333	28	13,4946	2,3214
PON_	516599	126	13,412	5,8175
ADJ_grande	30731	23	11,2075	4
ADV_presto	2152	9	11,0864	2
PRO:demo_questo	82344	36	10,3366	2,8333
VER:futu_tornare	1253	7	9,6252	0,4286
VER:futu_avere	4165	9	8,5928	1,8889
VER:futu_sbloccare	26	3	8,3906	1
VER:futu_potere	6253	10	8,2407	2
ADJ_assegnatole	2	2	7,8697	3,5
VER:futu_avvenire	316	4	7,1422	1,25
NOM_tifoso	5194	8	6,5972	2,5
PRO:poss_nostro	21368	14	6,4212	2,5714
VER:futu_giocare	1251	5	6,3238	4
VER:futu_aiutare	173	3	5,8827	2
NOM_soddisfazione	1592	5	5,8139	3,4
VER:futu_arrivare	1739	5	5,628	2
ADV_più	78862	26	5,4149	5,3462
VER:futu_piacere	25	2	5,3933	0,5
PRO:pers_ci	3535	6	5,3313	3,1667
VER:futu_sfruttare	27	2	5,3252	3,5
VER:futu_riprendere	272	3	5,2941	2
ADJ_Marek	28	2	5,293	2,5
VER:futu_riempire	40	2	4,9788	2
VER:futu_andare	2508	5	4,8642	0,4
VER:futu_regalare	67	2	4,5273	2
VER:infi_vincere	5017	6	4,4819	3,1667
VER:futu_lottare	85	2	4,3198	1
VER:futu_migliorare	93	2	4,2415	0
VER:futu_rivedere	97	2	4,2048	0,5
VER:futu_dimostrare	102	2	4,1611	0
ADV_insieme	8566	7	4,0948	2,7143
PRO:poss_suo	68316	21	4,0843	4,5714
VER:cpre_essere	15045	9	4,0576	3,8889
NOM_accordo	6009	6	4,0538	2,5
VER:futu_volere	749	3	3,9893	2,3333
NOM_bisogno	3904	5	3,9608	5,6
ADV_non	179623	40	3,9201	3,925
VER:futu_capire	149	2	3,8322	1
PRO:pers_lo	24369	11	3,749	2,6364
VER:futu_perdere	167	2	3,7334	2
NOM_elemento	2348	4	3,7325	2,5
VER:cpre_avere	4397	5	3,7229	4,2
NOM_	24724	11	3,696	3,9091
ADV_ancora	29108	12	3,6763	2,5
VER:futu_continuare	1093	3	3,5093	0,3333
NOM_traguardo	1097	3	3,5047	5,3333
VER:futu_raggiungere	225	2	3,4758	0,5
NOM_meglio	1162	3	3,432	4
NPR_Cesena	1208	3	3,3832	2,6667
PRO:refl_se	2957	4	3,356	0
ADJ_buon	3020	4	3,3218	3,25
VER:cpre_dovere	1302	3	3,289	2,3333
NOM_fiducia	3136	4	3,261	3,5
PRO:indef_tutto	19483	9	3,2439	3,2222
VER:futu_uscire	304	2	3,2165	0,5
NOM_altezza	1436	3	3,1663	2,6667
PRO:pers_la	5843	5	3,1651	2,6
NOM_squadra	24351	10	3,1496	3
PRO:pers_mi	20826	9	3,0426	3,6667
NOM_IN	1631	3	3,0076	5,6667
VER:infi_fare	34785	12	3,0046	4,0833

Dico che

Liste des cooccurrents spécifiques de *dico che* dans l'empan gauche de 4 items (tableau de gauche) et dans l'empan droit de 10 items (tableau de droite) dans le corpus PRES-IT :

Occ	Fréq	CoFréq	Indice	DistMoy
PRO:refl_vi	2517	5	7,18	0,60
ADV_quando	321	2	4,11	0,00

Occ	Fréq	CoFréq	Indice	DistMoy
VER:pres_intendere	338	3	5,23	4,67
VER:pper_ribadire	85	2	4,46	2,50
VER:pres_lavorare	160	2	3,91	3,00
NOM_impegno	3548	4	3,34	3,00

Liste des cooccurrents spécifiques de *dico che* dans l'empan gauche de 4 items (tableau de gauche) et dans l'empan droit de 10 items (tableau de droite) dans le corpus REP :

Occ	Fréq	CoFréq	Indice	DistMoy
PRO:pers_io	9564	61	108,02	0,08
PRO:refl_vi	2131	11	19,05	0,00
SENT_?	33142	21	17,08	1,48
SENT_.	1064900	102	12,67	1,54
ADV_non	179623	33	10,93	0,15
PRO:pers_ti	1097	6	10,83	0,00
CON_quando	16550	12	10,73	0,08
ADV_ecco	2676	5	6,81	1,00
CON_ma	94251	18	6,52	1,06
PON_-	148959	23	6,51	1,22
PRO:pers_gli	6822	6	6,15	0,00
PRO:demo_questo	82344	16	5,96	0,88
PON_"	516599	48	5,89	1,44
ADV_allora	4245	5	5,83	0,40
NOM_tifoso	5194	5	5,40	0,40
ADJ_aggiunto	2368	4	5,38	1,25
ADJ_realista	68	2	5,37	0,00
PRO:pers_mi	20826	8	5,32	0,38
PRO:demo_vi	235	2	4,29	0,00
ADV_no	4707	4	4,22	1,25
ADV_perciò	453	2	3,72	0,00
ADV_perché	6450	4	3,69	0,00
PON_:	183087	20	3,67	1,50
ADV_personalmente	495	2	3,65	0,50
NOM_ambizione	573	2	3,52	1,00
ADJ_strano	775	2	3,26	1,00
NOM_avversario	3754	3	3,20	1,00
VER:geru_essere	890	2	3,14	1,50
ADV_intanto	4004	3	3,12	0,33
PRO:poss_mio	16220	5	3,10	1,40
NOM_modulo	1043	2	3,01	0,50

Occ	Fréq	CoFréq	Indice	DistMoy
VER:pres_essere	385871	130	27,79	2,95
ADV_non	179623	69	17,86	3,25
VER:pres_dovere	19493	19	11,84	2,58
CON_ma	94251	37	10,20	5,27
PRO:refl_ci	37333	23	10,19	3,26
PRO:poss_mio	16220	14	8,25	2,64
ADV_meglio	6351	9	7,31	1,44
VER:infi_contemperare	6	2	6,76	4,50
PRO:pers_mi	20826	13	6,13	4,62
VER:pres_servire	3353	6	5,64	3,83
VER:impf_potere	3439	6	5,58	3,00
PRO:demo_quello	59301	21	5,43	4,76
CON_se	36768	16	5,36	3,38
VER:ppre_dirigere	286	3	5,32	4,00
VER:pres_bisognare	4281	6	5,04	0,83
ADV_bene	12342	9	4,96	4,44
VER:futu_fare	4461	6	4,94	2,50
VER:cpre_dovere	1302	4	4,84	0,75
VER:pres_stare	19838	11	4,82	3,91
NOM_calcio	7248	7	4,75	4,00
VER:pres_potere	26362	12	4,37	2,75
VER:infi_essere	31833	13	4,20	3,85
NOM_riguardo	763	3	4,06	4,33
PRO:pers_io	9564	7	4,00	1,86
VER:infi_arrivare	4180	5	3,97	3,60
CON_perché	20923	10	3,91	6,20
NOM_c'è	17430	9	3,82	2,67
NOM_onestà	184	2	3,71	3,00
ADJ_sicuro	2603	4	3,68	3,75
VER:cpre_essere	15045	8	3,54	4,13
ADJ_forte	8108	6	3,53	5,17
VER:cond_bisognare	244	2	3,47	0,00
NOM_filtro	250	2	3,45	2,50
PRO:pers_le	3160	4	3,37	1,25
VER:infi_giocare	5706	5	3,35	3,40
VER:pres_bastare	3292	4	3,30	1,00
VER:pper_dispiacere	301	2	3,29	1,00
PRO:pers_noi	13427	7	3,12	3,29
ADJ_bello	6827	5	3,01	4,20
NOM_bis	420	2	3,00	1,00

9.4 Script utilisé dans RStudio

```

1 #
2 #calculate expected frequencies based on what type of collocation being used
3 #
4
5
6 `calculate.expected.frequencies` <- function(df){
7
8 # df - dataframe - collocation containing pivot, col1, o11,o12,o21,o22 columns
9 #
10 df$R1 <- df$o11+df$o12
11 df$R2 <- df$o21+df$o22
12
13 df$C1 <- df$o11+df$o21
14 df$C2 <- df$o12+df$o22
15
16 df$N <- df$o11+df$o12+df$o21+df$o22
17
18 df$E11 <- df$R1*(df$C1/df$N)
19 df$E12 <- df$R1*(df$C2/df$N)
20
21 df$E21 <- df$R2*(df$C1/df$N)
22 df$E22 <- df$R2*(df$C2/df$N)
23
24
25 return (df)
26
27 }
28
29
30 `calculate.observed.frequencies` <- function(df,k=0){
31
32 # df - dataframe - collocation containing pivot, col1, o11, f1, f2, N, k columns
33 #
34 if(k!=0){
35 df$o11 <- df$f
36 df$o12 <- k*df$f1 - df$f
37 df$o21 <- df$f2 - df$f
38 df$o22 <- df$N-(k+1)*df$f1-df$f2+df$f
39 df$N <- df$N-df$f1
40
41 }else if("k" %in% names(df)){
42 df$k=as.numeric(df$k)
43 df$o11 <- df$f
44 df$o12 <- df$k*df$f1 - df$f
45 df$o21 <- df$f2 - df$f
46 df$o22 <- df$N-(df$k+1)*df$f1-df$f2+df$f

```

```

47 df$N <- df$N-df$f1
48 }else{
49 df$o11 <- df$f
50 df$o12 <- df$f1 - df$f
51 df$o21 <- df$f2 - df$f
52 df$o22 <- df$N-df$f1-df$f2+df$f
53
54 }
55
56 return (df)
57
58 }
59
60
61 `simple.collocation.measures` <- function(dfE){
62
63 #dfE <- calculate.expected.frequencies(df)
64 # t - collocation type
65 #
66 dfE$MI = log2(dfE$o11/dfE$E11)
67
68 dfE$localMI = dfE$o11 * log2(dfE$o11/dfE$E11)
69
70 dfE$simpleLL = 2*(dfE$o11 * log(dfE$o11/dfE$E11) - (dfE$o11-dfE$E11))
71 dfE$simpleLL = ifelse(dfE$E11 < dfE$o11, dfE$simpleLL, -dfE$simpleLL)
72
73 dfE$zScore = (dfE$o11-dfE$E11)/sqrt(dfE$E11)
74
75 dfE$tScore = (dfE$o11-dfE$E11)/sqrt(dfE$o11)
76
77
78 return (dfE)
79
80 }
81
82 `statistical.collocation.measures` <- function(dfE){
83
84 # dfE <- calculate.expected.frequencies(df)
85 #
86 # chi-square
87 dfE$chisq = (dfE$o11-dfE$E11)^2/dfE$E11 +
88 (dfE$o12-dfE$E12)^2/dfE$E12 +
89 (dfE$o21-dfE$E21)^2/dfE$E21 +
90 (dfE$o22-dfE$E22)^2/dfE$E22
91
92

```

```

92
93 df$chisq = ifelse(df$E11 < df$O11,df$chisq,-df$chisq)
94 # chi-squareCorr
95 #
96 df$chisqY = df$N *
97   (((
98     ((10^6)*abs((df$O11/10^3)*(df$O22/10^3) - (df$O12/10^3)*(df$O21/10^3)) - df$N/2)^2 /
99     df$R1)
100 / df$R2) / df$C1) / df$C2)
101
102 df$chisqY = ifelse(df$E11 < df$O11,df$chisqY,-df$chisqY)
103
104 # log-likelihood
105 df$log_l = 2 * (
106   ifelse(df$O11>0, df$O11*log(df$O11/df$E11), 0) +
107   ifelse(df$O12>0, df$O12*log(df$O12/df$E12), 0) +
108   ifelse(df$O21>0, df$O21*log(df$O21/df$E21), 0) +
109   ifelse(df$O22>0, df$O22*log(df$O22/df$E22), 0)
110 )
111 df$log_l = ifelse(df$E11 < df$O11,df$log_l,-df$log_l)
112
113 #average-MI
114 df$avgMI = df$O11*log2(df$O11/df$E11)+
115           df$O12*log2(df$O12/df$E12) +
116           df$O21*log2(df$O21/df$E21) +
117           df$O22*log2(df$O22/df$E22)
118
119 #dice
120 df$Dice = 2 * df$O11 / (df$R1 + df$C1)
121
122
123 #odds-ratio
124 df$log.odds = log(
125   ((df$O11 + .5) * (df$O22 + .5)) /
126   ((df$O12 + .5) * (df$O21 + .5))
127 )
128
129 return (df)
130 |
131 }
132
133
134 calculate.collocation.signature' <- function(dF, k=0,outFile="collocationResult.tsv"){
135
136 # dF - name or path of the file containing data
137 # tab separated value must containe f, f1,f2, N columns

```

```

138 # or data-frame containing same columns
139
140 if(is.data.frame(dF)){
141   df=dF
142 }else{
143   df<- read.delim(dF)
144 }
145
146 if(!("f" %in% names(df))||
147    !("f1" %in% names(df))||
148    !("f2" %in% names(df))||
149    !("N" %in% names(df))){
150   stop("Data must have following columns: f,f1,f2,N")
151 }else{
152   df$f = as.numeric(df$f)
153   df$f1 = as.numeric(df$f1)
154   df$f2 = as.numeric(df$f2)
155   df$N = as.numeric(df$N)
156
157   df <- calculate.observed.frequencies(df,k)
158   df <- calculate.collocation(df)
159   write.table(df, file=outFile, sep="\t")
160   return(df)
161 }
162 }
163
164 }
165 calculate.collocation.observed' <- function(dF,outFile="collocationResult.tsv"){
166
167 if(is.data.frame(dF)){
168   df=dF
169 }else{
170   df<- read.delim(dF)
171 }
172
173 if(!("O11" %in% names(df))||
174    !("O21" %in% names(df))||
175    !("O21" %in% names(df))||
176    !("O22" %in% names(df))){
177
178   stop("Data must have following columns: O11,O12,O21,O22")
179 }else{
180
181
182
183   df <- calculate.collocation(df)

```

```

184     write.table (df, file=outFile, sep="\t")
185     return(df)
186
187   }
188 }
189
190 }
191
192 `calculate.collocation` <- function(df){
193
194   df <- calculate.expected.frequencies(df)
195   df <- simple.collocation.measures(df)
196   df <- statistical.collocation.measures(df)
197   return(df)
198 }
199 }
200
201
202
203
204 `calculate.frequencies` <- function(...){
205
206 #recalculates the sample size in the case where collocates or/and pivot are n-grams
207 #arguments: N - original sample size
208 #           f1 - pivot frequency
209 #           n1 - no. tokens in the pivot
210 #           f2 - collocate frequency
211 #           n2 - no.tokens in the collocate
212 # @result - df containing the input data with the corrected size in the Nc column
213 #
214 #- data can be passed as tsv file, data frame or as vectors
215 #Examples: calculate.frequencies(file.name="nameOfTheDataFile")
216 #          calculate.frequencies(df=dataFrame)
217 #          calculate.frequencies(N=1000, f1=20, n1=2)
218 #          calculate.frequencies(N=1000, f2=20, n2=2)
219 #          calculate.frequencies(N=1000, f1=40, n1=3, f2=20, n2=2)
220 #
221 #
222 arguments <- list(...)
223
224 if("file.name" %in% names(arguments)){
225   CFdf <- read.delim(arguments$file.name)
226
227 }else if("df" %in% names(arguments)){
228
229

```

130:3 `statistical.collocation.measures`(df) ↕

R Script ↕

```

230   CFdf <- arguments$df
231
232 }else{
233   CFdf <- data.frame(arguments[1])
234   if(length(names(arguments))>=2){
235     for(i in 2:length(names(arguments))){
236       CFdf[names(arguments)[i]] <- arguments[i]
237     }
238   }
239 }
240 }
241
242 if(!"N" %in% colnames(CFdf)){
243   stop("original size N missing from the input data ")
244 }
245
246 #fill in missing frequencies
247 if(!"f1" %in% colnames(CFdf)){
248   CFdf$f1 <- 0
249 }
250 if(!"n1" %in% colnames(CFdf)){
251   CFdf$n1 <- 1
252 }
253 if(!"f2" %in% colnames(CFdf)){
254   CFdf$f2 <- 0
255 }
256 if(!"n2" %in% colnames(CFdf)){
257   CFdf$n2 <- 1
258 }
259
260 CFdf$Nc <- CFdf$N - (CFdf$n1-1)*CFdf$f1 - (CFdf$n2-1)*CFdf$f2
261
262
263   return(CFdf)
264 }
265 }
266
267
268
269

```

130:3 `statistical.collocation.measures`(df) ↕

R Script ↕

