

## Comment différencier des adverbes de posture épistémique ? Proposition d'une méthodologie

*How to differentiate epistemic stance adverbs? Proposal of a methodology*

Corinne Rossari et Laura Aubry

---



### Édition électronique

URL : <https://journals.openedition.org/praxematique/9281>

ISSN : 2111-5044

### Éditeur

Presses universitaires de la Méditerranée

Ce document vous est fourni par Université de Neuchâtel



### Référence électronique

Corinne Rossari et Laura Aubry, « Comment différencier des adverbes de posture épistémique ? Proposition d'une méthodologie », *Cahiers de praxématique* [En ligne], 82 | 2024, mis en ligne le 01 octobre 2024, consulté le 07 octobre 2024. URL : <http://journals.openedition.org/praxematique/9281>

---

Ce document a été généré automatiquement le 4 octobre 2024.



Le texte seul est utilisable sous licence CC BY-NC-ND 4.0. Les autres éléments (illustrations, fichiers annexes importés) sont « Tous droits réservés », sauf mention contraire.

---

# Comment différencier des adverbes de posture épistémique ?

## Proposition d'une méthodologie

*How to differentiate epistemic stance adverbs? Proposal of a methodology*

Corinne Rossari et Laura Aubry

---

### 1 Introduction

- 2 Le but de notre article<sup>1</sup> est de proposer une méthodologie qui permette de conjuguer étroitement une analyse qualitative conçue dans le cadre de la linguistique énonciative avec une étude empirique fondée sur l'utilisation des formes prises en compte par les locuteur-trices<sup>2</sup> et une étude textométrique faisant apparaître les environnements linguistiques privilégiés de ces formes dans des corpus. Nous fonderons notre étude sur quatre formes qui appartiennent à la catégorie des adverbes de posture épistémique : *certainement*, *probablement* et *peut-être*. Notre sélection est motivée par un éventail couvrant les différentes forces associées traditionnellement à ces adverbes en termes de 'certitude', et par le fait de devoir limiter le paradigme, compte tenu de la méthodologie, pour prendre en compte, en sus de l'analyse qualitative de ces formes, celle de leur usage effectif, par le biais de deux études : celle concernant la manière dont les locuteur-trices les utilisent dans de courts énoncés conçus par nos soins et celle de leur usage dans des corpus, étude nécessitant le traitement de grandes quantités de données.
- 3 Nous commencerons par un survol des classifications de ces adverbes dans des ouvrages de référence, classifications qui les différencient en fonction d'une force modale (1.1.). Après avoir mis en relief les limites quant à la possibilité de différencier ces formes dans ce cadre, nous en proposerons une description énonciative fondée sur le type de prise en charge des contenus qu'ils prennent dans leur portée (1.2.). Pour établir la nature de cette prise en charge, nous nous adosserons à la théorie de la polyphonie linguistique élaborée par Ducrot (1984, 1989, 2001). Au moyen de ces outils, un profil sémantique (PS) pourra être défini pour chacune des formes, permettant de

relever ce qui les singularise. Ce PS sera ensuite mobilisé pour faire des hypothèses sur leur potentiel d'utilisation (PU), qui sera testé avec différentes méthodes empiriques. Nous commencerons par une méthode fondée sur un test d'utilisation effective de ces marqueurs par des locuteur·trices francophones (2.1.). Puis nous utiliserons des calculs statistiques pour voir dans quels genres de discours ces adverbes apparaissent de façon statistiquement significative (2.2.), et dans un troisième temps, nous identifierons les environnements linguistiques privilégiés de ces adverbes (2.3.). Enfin, pour mieux visualiser la proximité ou la distance entre ces quatre adverbes, nous recourrons à une analyse factorielle des correspondances fondée sur des co-fréquences dans un corpus constitué par des extraits de textes appartenant à des genres discursifs différents (2.4.).

- 4 Nous avons sélectionné des corpus représentant trois genres discursifs variant au niveau de la dimension argumentative : le genre encyclopédique, le genre presse, et le genre discours politique<sup>3</sup>. Ces trois genres permettent de faire apparaître ou de mettre en retrait la dimension argumentative dont un discours peut être plus ou moins investi en fonction de son objectif. Le discours encyclopédique est un discours ayant vocation d'informer ; en cela, la dimension argumentative n'a pas à être mise en avant. À l'inverse, le but du discours politique est avant tout de convaincre, avant même d'informer ; sa dimension argumentative est donc constitutive du genre. À mi-chemin nous avons sélectionné un discours beaucoup plus hétérogène en termes de genre, le discours de presse, dont les parties font plus ou moins ressortir la dimension argumentative : certaines parties, comme les articles d'opinion, sont porteuses de cette dimension, alors que d'autres, comme les brèves, la revue de presse, ne mettent *a priori* pas en avant une telle dimension. La prise en compte de ce genre permet ainsi d'avoir un cas intermédiaire entre un genre à faible dimension argumentative et un genre à forte tendance argumentative.

## 1. Analyse qualitative

### 1.1. Approche liée à la force modale

- 5 Les adverbes *certes*, *certainement*, *probablement* et *peut-être* relèvent du paradigme des adverbes modaux dans les grammaires linguistiques comme la *Grande Grammaire du français* (Abeillé & Godard, 2021) ou la *Grammaire critique du français* (Wilmet, 2003). Leur sémantique est considérée comme manifestant une oscillation entre le possible et le certain. Dans la *Grande Grammaire du français* (GGF), ils sont décrits comme permettant de qualifier la valeur de vérité d'une proposition (2021 : 927) – à noter que *certes* ne figure ni dans les adverbes modaux ni dans d'autres catégories d'adverbes dans cette grammaire. Dans la *Grammaire critique du français*, ils sont qualifiés de « modulateurs de vérité », accompagnés de : *à coup sûr*, *pour sûr*, *pour ça* [...], et *certes* fait partie de cette liste (2003 : 576). Si on approfondit les sous-catégories d'adverbes modaux envisagées dans la GGF (2021 : 927), on en relève cinq comportant les formes suivantes :

Nécessité :

*fatalement*, *forcément*, *immanquablement*, *inéluçtablement*, *inévitablement*, *infailliblement*, *nécessairement*, *obligatoirement*

Certitude :

*assurément*, *bien entendu*, *bien sûr*, *certainement*, *incontestablement*, *indéniablement*,

*indiscutablement, indubitablement, sans conteste, sans aucun/nul doute, sûrement*

Possibilité ou plausibilité – épistémique :

*peut-être, plausiblement, possiblement, probablement, supposément, vraisemblablement, sans doute*

Visibilité (liée à un adjectif) :

*apparemment, évidemment, manifestement, visiblement* (proposition présentée comme vraie ou certaine sur la base de son degré d'évidence).

Groupe résiduel :

*effectivement, naturellement*

- 6 *Certainement, probablement* et *peut-être* sont distribués entre les deux catégories 'certitude' et 'plausibilité épistémique'. Ces deux catégories représentent deux niveaux d'une échelle conçue en termes de degrés de certitude par rapport à la vérité de l'énoncé : *peut-être* et *probablement* sont dans la même catégorie et relèvent d'un degré plus faible par rapport à *certainement*, qui est dans une autre catégorie, associée à un degré de certitude plus fort.
- 7 Dans les travaux linguistiques ayant trait à la classification des adverbes, ceux-ci sont envisagés en fonction du type de certitude qu'ils communiquent et ils sont souvent situés sur une échelle exprimant différents degrés de certitude ou force modale. Dans Molinier & Levrier (2000 : 92), *peut-être, probablement, certainement* et *certes* sont classés dans la catégorie des adverbes modaux. Les auteurs distinguent quatre regroupements sémantiques fondés sur la modalité exprimée par l'adjectif sur lequel les adverbes sont formés ('modalité aléthique du nécessaire', 'modalité épistémique du certain', 'modalité aléthique du possible' ou 'modalité épistémique du plausible' et les adverbes exprimant la notion de visibilité) (2000 : 105). *Probablement* et *certainement* sont décrits comme relevant de la plausibilité – l'adjectif *certain* « subit un affaiblissement de sens lorsqu'il devient la base de l'adverbe » (2000 : 94), mais *certainement* peut exprimer la certitude lorsqu'il est employé seul en réponse à une question totale ou dans la construction [Adv, que P!] (2000 : 94-95). Chez De Cesare (2022 : 523-524) ils sont classés selon un degré de force épistémique dans l'ordre suivant : *peut-être, probablement* et *certainement* sur le dernier échelon, *certes* n'étant pas pris en compte dans cette classification. On retrouve un ordre identique pour les trois adverbes quant à leur disposition dans Gosselin (2010 : 86), qui envisage une échelle allant de la « certitude que non p » à la « certitude que p ». L'étude de Vold (2008 : 65) conçoit une échelle qui intègre les expressions épistémiques dans leur ensemble, où figurent adverbes et constructions à sens modal épistémique. On y retrouve *peut-être, probablement* et *certainement*, disposés en ordre croissant de certitude. Si *certes* n'apparaît pas dans ces travaux, il est pris en compte dans d'autres études aux côtés de ces trois adverbes : par exemple, Borillo (1976 : 87-88) le classe parmi les adverbes modalisateurs de l'assertion, Gil (1995 : 421-422) le considère comme un marqueur de certitude ('Gewissheit'), au même titre que *certainement*. Toutefois, *certes* est également mis en relation avec la prise en charge dans des études faites dans une perspective énonciative : selon Anscombe (1981 : 117-118), « *certes* dériverait un acte de concession d'un renforcement de l'assertion ». Son analyse est fondée sur deux usages de *certes* : *certes* de renforcement (similaire à *certainement*), et *certes* concessif. Adam (1997 : 9) reprend son analyse en ajoutant une composante liée

au degré de vérité : « Tout dépend, en fait, du degré de vérité accordé par le locuteur au propos qu'il attribue polyphoniquement à E1 : moins l'adhésion du locuteur est forte, plus l'interprétation concessive devient plausible. »

- 8 La classification en termes de force modale se retrouve aussi dans la littérature anglophone. Nuyts (2014: 38) affirme que : « *epistemic modality may be considered a matter of degree, i. e., as involving a scale, at least going from absolute certainty via probability to fairly neutral possibility that the state of affairs is real* ». Les grammaires de l'anglais telles que la *Comprehensive Grammar of the English Language* (Quirk et al., 1985), *The Longman Grammar of Spoken and Written English* (Biber et al., 1999), et la *Cambridge Grammar of the English Language* (Huddleston & Pullum, 2002) classifient également les adverbes épistémiques « en fonction de leur force modale » (Simon-Vandenberghe & Aijmer, 2007 : 63)<sup>4</sup>.
- 9 Les classifications fondées sur une échelle de type *force modale* posent question quant à la possibilité de discriminer ces adverbes. Deux problèmes se posent : d'une part le statut de *certes*, qui, dans plusieurs études, est considéré comme donnant également une indication de certitude et ce, même dans les descriptions qui lui sont consacrées spécifiquement (cf. Anscombre, 1981 et Adam, 1997) ; d'autre part, le statut de faible certitude associé à *peut-être*, qui se trouve toujours en bas de l'échelon représentant différentes forces modales.
- 10 Pour tester la discrimination de ces adverbes en fonction d'échelles, on peut prendre en compte deux types de graduation : une graduation qui va du pôle 'faible certitude' au pôle 'forte certitude' dans une représentation où l'état de choses sous-jacent à la proposition reste non actualisé, et une autre qui s'applique directement à l'état de choses même, et qui va du pôle 'désactualisé' au pôle 'actualisé'.

#### Échelle 1



- 11 La première échelle ne convient pas dans la mesure où *certes*, à la différence de *certainement*, conduit à une interprétation où la force de la certitude n'est pas pertinente – l'état de choses sur lequel il porte étant compris comme actualisé. *Certes Paul est arrivé* suppose qu'il est attesté que Paul est arrivé, mais *Certainement, Paul est arrivé*, suppose que son arrivée n'est pas attestée.

#### Échelle 2



- 12 La seconde échelle semblerait plus conforme à une disposition graduelle de ces quatre adverbes avec *certes* occupant la dernière place. Toutefois, le test au moyen de *même* utilisé par Ducrot (1980) pour situer les arguments selon leur force sur une même échelle argumentative ne fonctionne pas. Le test préconise que si deux propositions (*p* et *p'*) peuvent mener à une même conclusion *r* sur une échelle donnée, une phrase comme *p* ou (*et*) *même p'* présuppose que *p'* est supérieur à *p* sur cette échelle (1980 : 18).

Or, ce test ne permet pas de montrer que ces adverbes sont différenciables en termes de force. Premièrement, *certes* semble ne pas pouvoir se situer sur ce type d'échelle :

[1] ??Il est certainement arrivé, il est même certes arrivé.

- 13 Deuxièmement, *peut-être* n'est, dans certains emplois en tout cas, pas situable sur un continuum par rapport aux autres adverbes. Si, dans un emploi comme le suivant, il semble indiquer un degré moindre de certitude :

[2] J'ai peut-être oublié mes clés, j'ai même probablement oublié mes clés, il faut revenir à la maison.

- 14 *peut-être* a aussi la propriété de pouvoir être utilisé avec des énoncés exprimant des vérités générales ou des adages :

[3] L'habit ne fait peut-être pas le moine.

- 15 Dans ce type d'énoncé *peut-être* ne situe pas le contenu sur une échelle de certitude, il est difficilement combinable avec un autre énoncé modalisé par *probablement* ou *certainement*, qui serait présenté par *même* comme un argument plus fort en vue d'une conclusion :

[4] ??L'habit ne fait peut-être pas le moine et même l'habit ne fait probablement/certainement pas le moine, il faut se méfier des apparences.

- 16 Si *peut-être* est utilisable dans ce type d'énoncé, c'est pour permettre une prise de distance du locuteur par rapport à l'adage :

[5] L'habit ne fait peut-être pas le moine, mais en entrevue, notre apparence physique doit refléter le poste que nous cherchons à avoir. Soignez donc votre présentation physique et votre parler. <<https://www.atoutrecrutement.com/recherche-dun-emploi-cadre-les-cles-de-la-reussite/>>

- 17 *Probablement* ou *certainement* seraient sentis comme très peu naturels dans le même contexte, à la différence de *certes*, qui comme *peut-être* permet une prise de distance :

[6] ??L'habit ne fait probablement/certainement pas le moine, mais en entrevue, notre apparence physique doit refléter le poste que nous cherchons à avoir. Soignez donc votre présentation physique et votre parler.

[7] L'habit ne fait certes pas le moine, mais en entrevue, notre apparence physique doit refléter le poste que nous cherchons à avoir. Soignez donc votre présentation physique et votre parler.

- 18 Ces quelques manipulations d'exemples montrent qu'une description en termes de degrés de certitude ne rend pas compte des propriétés combinatoires de ces adverbes. Nous allons alors explorer une autre façon de les caractériser, en nous interrogeant sur le type de prise en charge qui peut être associé aux énoncés dans lesquels ils figurent.

## 1.2. Approche liée à l'énonciation

- 19 Pour isoler les propriétés énonciatives de ces adverbes, nous faisons un bref rappel de la théorie de la polyphonie linguistique mise au point par Ducrot (1984, 1989) et Ducrot (2001), qui porte spécifiquement sur l'utilité de la notion d'«énonciateur». Nous nous

référons surtout à Ducrot (2001), qui explicite le type de rapport que le locuteur (au sens de 'responsable d'un contenu') peut tisser avec l'énonciateur. Dans cette théorie, l'énoncé est analysé comme mettant en scène différents énonciateurs, eux-mêmes sources de points de vue, envers lesquels le locuteur peut manifester différentes attitudes : il peut « les donner comme ses porte-parole » (Ducrot, 2001 : 2), cas où le locuteur « reprend le point de vue à son propre compte » (2001 : 6), ce qui correspond à une relation de *prise en charge* ; « leur donner son accord » (2001 : 2), ce que nous désignons par la relation d'*accord* ; ou « s'opposer à eux » (2001 : 2), ce qui correspond à une relation de *refus* ou *désaccord*.

- 20 Au lieu de voir les adverbes comme se différenciant au niveau d'une force modale, nous faisons l'hypothèse qu'ils donnent des indications sur la relation que le locuteur tisse avec l'énonciateur auquel un point de vue est attribué. Ces indications conditionnent leur possibilité d'emploi, notamment le fait que *peut-être* ou *certes* peuvent être utilisés en vue de remettre en cause la prise en charge d'un contenu, alors que *certainement* ou *probablement* sont moins facilement utilisables dans ce type de contexte (cf. ex [6]).
- 21 En utilisant ces différents types de relations en lien avec le point de vue concerné par l'adverbe, chaque adverbe est caractérisable ainsi :
- (i) *Probablement* et *certainement* indiquent que le locuteur tisse une relation de 'prise en charge' avec l'énonciateur auquel un point de vue est attribué. Ce point de vue concerne le contenu.
  - (ii) *Certes* indique que le locuteur tisse une relation d'« accord' avec l'énonciateur auquel un point de vue est attribué. Ce point de vue concerne le contenu, comme indiqué dans la description de Ducrot (2001 : 5).
  - (iii) *Peut-être* indique que le locuteur tisse une relation d'« accord' avec l'énonciateur auquel un point de vue est attribué. Ce point de vue concerne le contenu.
  - (iv) *Peut-être* indique que le locuteur tisse une relation d'« accord' avec l'énonciateur auquel un point de vue est attribué. Ce point de vue concerne l'énonciation.
- 22 Au moyen de ces outils nous différencions donc *probablement* et *certainement* de *certes* et *peut-être* par rapport à la nature de la relation ('prise en charge' et 'accord'), et *certes* de *peut-être* par rapport au point de vue : *peut-être* peut agir sur un point de vue qui concerne le contenu ou l'énonciation, alors que *certes* agit sur un point de vue qui ne concerne que le contenu.
- 23 Ce profil sémantique rend compte du fait que *certes* et *peut-être* permettent aisément des enchaînements en *mais*, étant donné que la relation d'« accord' qui les caractérise est propice à la concession. Avec un enchaînement au moyen de *mais*, la prise en charge des deux contenus articulés ne peut pas être du même ordre : *mais* indique que le contenu sur lequel il enchaîne est *accordé* et que celui qui le suit est *pris en charge*. *Peut-être* peut également permettre un enchaînement en *mais* conduisant à une remise en cause du contenu, ce que *certes* ne peut pas faire. Ce fonctionnement peut ressortir avec la construction de l'exemple suivant :
- [8] Paul a peut-être cessé de fumer, mais je l'ai vu griller une cigarette hier.  
[9] ??Paul a certes cessé de fumer, mais je l'ai vu griller une cigarette hier.
- 24 Si ce type d'enchaînement est constructible avec *peut-être*, c'est que cet adverbe peut prendre dans sa portée non pas un contenu, mais une énonciation. C'est cette dernière qui est la cible de l'adverbe. Cela permet un enchaînement qui s'émancipe en quelque sorte du contenu : peu importe ce qui est mentionné dans ce contenu, *peut-être* indique

qu'il accorde l'énonciation, ce qui permet au locuteur de rebondir sur quelque chose qu'il estime non pertinent. Cet emploi de *peut-être* a été décrit dans Rossari *et al.* (à paraître) comme post-modal. *Certes* est incompatible dans un tel contexte, car il porte sur le contenu en indiquant que le locuteur l'accorde. Il ne peut donc pas être suivi d'un énoncé qui rend ce contenu non pertinent.

- 25 *Probablement* et *certainement*, du fait qu'ils indiquent une relation de prise en charge, sont moins attendus dans une structure concessive (cf. [6]) et *a fortiori* quand celle-ci permet la remise en cause du contenu qui précède :

[10] ??Paul a probablement/certainement cessé de fumer, mais je l'ai vu griller une cigarette hier.

- 26 Dans ce type d'emploi, le locuteur a l'air de se contredire ; il reprend le point de vue à son compte tout en le faisant suivre d'un contenu qui le remet en cause. En revanche, un enchaînement en *donc* qui signale que le locuteur introduit une conclusion suite à un argument qu'il prend en charge est adéquat :

[11] Paul a probablement/certainement cessé de fumer récemment, il va donc être un peu nerveux.

- 27 En synthèse le profil sémantique de chaque adverbe est le suivant :

PS de *probablement* et *certainement* : le contenu de l'énoncé dans lequel l'adverbe intervient est pris en charge, ce qui met ce contenu sur le devant de la scène discursive ;

PS de *certes* : le contenu de l'énoncé dans lequel l'adverbe intervient est accordé, ce qui le met à l'arrière-plan de la scène discursive ;

PS de *peut-être* : le contenu de l'énoncé est accordé (même fonctionnement que *certes*) ; l'énonciation du contenu est accordée (emploi particulier à *peut-être*), ce qui conduit ce contenu à être mis entre parenthèses (le locuteur le rejette comme non pertinent).

- 28 Nous allons passer au volet empirique de cette étude en testant le potentiel d'utilisation qui peut être issu de chacun de ces profils sémantiques.

## 2. Analyse empirique

- 29 Le profil sémantique des adverbes donne des indications sur le potentiel d'utilisation de chacun d'eux.

- 30 Nous allons vérifier ce potentiel dans un premier temps au moyen d'un test sur les préférences des locuteur-trices francophones concernant le choix d'un adverbe dans une configuration adaptée aux propriétés attribuées à chacun d'entre eux. Dans un second temps, dans une perspective qui associe étroitement une démarche *corpus-based* et *corpus-driven*<sup>5</sup>, nous allons mobiliser une étude utilisant des calculs et des méthodes statistiques pour identifier d'une part, leur tendance à être sur ou sous-utilisés selon des corpus qui varient en termes de genres, et d'autre part, selon les environnements linguistiques auxquels ils sont associés de façon significative.

## 2.1. Test statistique sur le choix des locuteur·trices

- 31 Nous avons soumis à 49 sujets francophones trois énoncés en leur demandant de choisir parmi les quatre adverbes, avec la possibilité de sélectionner deux adverbes par énoncé au maximum. Les énoncés [12], [13] et [14] ont été conçus pour s'adapter aux caractéristiques du profil sémantique de chaque adverbe.

[12] Je suis [ADVmodal] stupide, mais j'ai tout compris.

[13] Paul est [ADVmodal] arrivé, mais en retard.

[14] Les enfants dorment [ADVmodal], donc ils ne remarqueront rien.

- 32 L'énoncé [12] a été conçu pour rendre tangible la possibilité de remettre en cause le contenu de l'énoncé où figure l'adverbe avec la suite en *mais*. Cette caractéristique est en phase avec le PU de *peut-être* utilisable dans ce type de séquence par le biais de son profil, qui lui permet d'agir sur l'énonciation du contenu dans lequel il apparaît. Ce contenu apparaît alors comme non pertinent, ce qui rend possible sa remise en cause. L'adverbe attendu prioritairement est *peut-être*.
- 33 L'énoncé [13] a été conçu pour illustrer la configuration de la concession, avec un contenu accordé, qui est mis à l'arrière-plan du discours en cours, celui dans lequel l'adverbe intervient, et un contenu présenté comme pris en charge. Cette configuration est propice à l'emploi de *peut-être* et de *certes* qui ont la propriété de signaler une relation d' 'accord' concernant le contenu d'un point de vue. Les adverbes attendus prioritairement sont *certes* et *peut-être*.
- 34 L'énoncé [14] a été conçu pour illustrer la configuration de l'enchaînement argumentatif dans lequel l'énoncé qui sert d'argument est pris en charge par le même locuteur que celui qui sert de conclusion. Cette configuration est propice à l'emploi de *probablement* ou *certainement*, qui tous deux ont la propriété de signaler une relation de 'prise en charge' concernant le contenu d'un point de vue. Les adverbes attendus prioritairement sont *probablement* et *certainement*.

Tableau 1 : Choix des locuteur·trices

	1	2	3	4	(1, 2)	(1, 3)	(1, 4)	(2, 3)	(2, 4)	(3, 4)
Ph1	3	22	0	2	16	0	1	1	2	2
Ph2	22	5	5	4	4	6	0	1	0	1
Ph3	0	1	22	11	0	0	0	1	0	14

1 : certes

2 : peut-être

3 : certainement

4 : probablement

Ph1 : Je suis [...] stupide mais j'ai tout compris.

Ph2 : Paul est [...] arrivé, mais en retard.

Ph3 : Les enfants dorment [...] donc ils ne remarqueront rien.

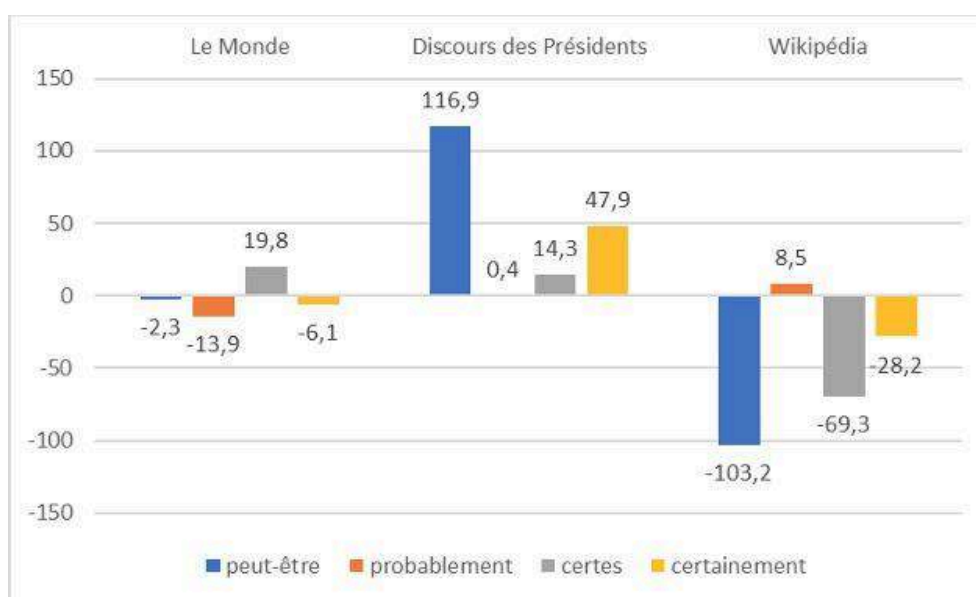
- 35 Le résultat du test du Khi-deux<sup>6</sup> permet de rejeter l'indépendance des variables (choix aléatoire des adverbes dans ces trois exemples) avec une confiance supérieure à 99,5 %.

- 36 Les adverbes sont choisis prioritairement en lien avec ce qui était attendu par la constitution de l'exemple. Toutes les données recueillies dans le tableau 1<sup>7</sup> confirment que l'utilisation de l'adverbe répond donc au potentiel d'utilisation qui peut être tiré du profil sémantique que l'analyse qualitative lui prête.

## 2.2. Test statistique concernant les genres discursifs dans lesquels les adverbes apparaissent de façon significative

- 37 Les caractéristiques issues du profil sémantique de chaque adverbe donnent des indications sur leur potentiel d'utilisation dans des corpus. Ce PU peut être testé par une étude quantitative des genres dans lesquels ces formes sont le plus utilisées.
- 38 Nous avons pris en compte trois genres qui font varier la dimension argumentative des discours : pour rappel, le genre encyclopédique, qui gomme cette dimension, le genre presse, qui est intermédiaire (il ne la gomme pas, mais ne la revendique pas non plus) et le genre discours politique, genre dans lequel la dimension argumentative est constitutive du discours. Les formes liées à l'« accord » peuvent évoquer une mise en contraste de points de vue, mise en contraste typique des discours à forte dimension argumentative. Le PU de *certes* et *peut-être* les rendrait donc propices à être utilisés dans des discours à forte dimension argumentative, ce qui les distinguerait de *probablement* et *certainement*.
- 39 Pour tester cette hypothèse, nous avons créé un corpus composé de trois parties différentes représentant les trois genres discursifs mentionnés plus haut, afin de faire des calculs statistiques sur la « représentativité » des formes en cause. Le calcul de représentativité est fondé sur l'écart par rapport à la distribution attendue dans chacune des parties du corpus total (intégrant les trois genres de discours). La distribution attendue est celle où les adverbes se répartissent de façon équitable compte tenu de leur nombre d'occurrences dans le corpus total.
- 40 La figure 1 ci-dessous représente les indices de spécificité calculés par la plateforme TXM (Heiden *et al.*, 2010). Le seuil de banalité est entre  $-2$  et  $+2$  ; les valeurs supérieures peuvent être considérées comme une sur-représentation, les inférieures comme une sous-représentation<sup>8</sup>.

Figure 1 : Indices des spécificités des adverbes par partie



- 41 La différence nette qui ressort concerne *peut-être* et *certes*. Les deux adverbes sont nettement sur-représentés dans le corpus *Discours des Présidents* (à dimension fortement argumentative) et nettement sous-représentés dans le corpus encyclopédique. Ces indices sont en phase avec leur PU, qui prévoit une composante dialogique dans leurs emplois, composante que l'on retrouve davantage dans des discours à forte dimension argumentative. Le corpus de presse manifeste des indices de représentation moins contrastés entre les adverbes : seul *certes* reste sur-représenté, les trois autres sont sous-représentés avec des indices nettement moins élevés que dans le discours encyclopédique. De façon générale, *peut-être*, *certes* et *certainement* suivent les tendances des marqueurs de subjectivité analysés par Pamuksaç & Sanvido (2023) : ils sont sur-représentés dans le genre où la dimension argumentative est la plus marquée. *Probablement* s'écarte de ce comportement en étant sur-représenté dans le discours dont la dimension argumentative est la moins marquée.

### 2.3. Étude des environnements linguistiques dans lesquels les adverbes apparaissent de façon statistiquement significative

- 42 Le profil sémantique de ces adverbes devrait également conduire à des différences quant aux environnements linguistiques spécifiques qui leur sont associés. Par « environnements spécifiques », nous entendons tous les cooccurrents accompagnant de façon statistiquement significative l'adverbe dans un empan pertinent. Pour identifier ces cooccurrents, nous utilisons l'indice de TXM ; les valeurs égales ou supérieures à 3 indiquent que la cooccurrence a plus de 99,9 % de chances de ne pas être due au hasard dans l'empan donné (cf. Leblanc *et al.*, 2017 : 141). L'empan de 11 *tokens* avant et après le mot pivot qui est l'adverbe (ADV) est celui qui présente la meilleure proportion d'emplois où l'adverbe est lié sémantiquement aux formes qui sont ses cooccurrents spécifiques<sup>9</sup>.
- 43 À partir de l'extraction de l'ensemble de ces cooccurrents ont été retenus ceux qui permettaient de souligner les traits associés à ces adverbes à la suite de l'étude de leur

profil sémantique. Nous avons ainsi relevé quatre catégories d'expressions en lien avec une forme ou une autre de polyphonie ou en lien avec des indications argumentatives. Ces catégories devraient singulariser *peut-être* et *certes* par rapport aux deux autres adverbes : les marques de la polyphonie répondent au trait « accord », qui les caractérise tous deux, et celles de la dimension argumentative font écho à la prédisposition de ces adverbes à intervenir dans des discours à forte dimension argumentative (comme relevé dans la section précédente).

44 Pour les marqueurs de polyphonie, les formes sélectionnées relèvent de trois catégories :

1. Les indications qui signalent un potentiel dialogique fictif ou mis en scène, avec des formes telles que la négation en *pas*, *rien*, *jamais*, indices d'une polyphonie selon Ducrot (1984 : 217-218) ; le point d'interrogation et d'exclamation, indices de contenus *adressés* au sens de Carel (2010) à un auditeur fictif ou réel<sup>10</sup> ; la forme *pourquoi* dans la mesure où elle permet de (re)construire une demande adressée ou non (*je me demande pourquoi / Pourquoi ?*).

2. Les indices de la présence du locuteur au moyen des pronoms de première personne ont également été considérés comme des indications propices au dialogue dans la mesure où la mise en scène du locuteur à la première personne peut donner à l'énoncé un caractère *adressé*.

3. Une autre catégorie prise en compte regroupe les verbes de dire, du fait qu'ils permettent d'une manière ou d'une autre de faire entendre une voix (*dire*, *affirmer*, *justifier...*).

45 Dans la catégorie 'indications argumentatives', nous avons relevé :

4. Celles qui de manière conventionnelle servent à orienter l'énoncé vers une certaine conclusion ou à renforcer/amoinrir la force argumentative d'un énoncé, formes qui relèvent chez Ducrot de la catégorie des « opérateurs argumentatifs » (cf. Ducrot, 1983 : 9-10) telles que *trop*, *peu*, *mieux*, *même*, *bien*, *beaucoup*, *très...*

### 2.3.1. Résultats pour la famille 1 (polyphonie dialogique)

Figure 2 (à gauche) : Polyphonie type dialogique (*peut-être*)

Figure 3 (à droite) : Polyphonie type dialogique (*probablement*)

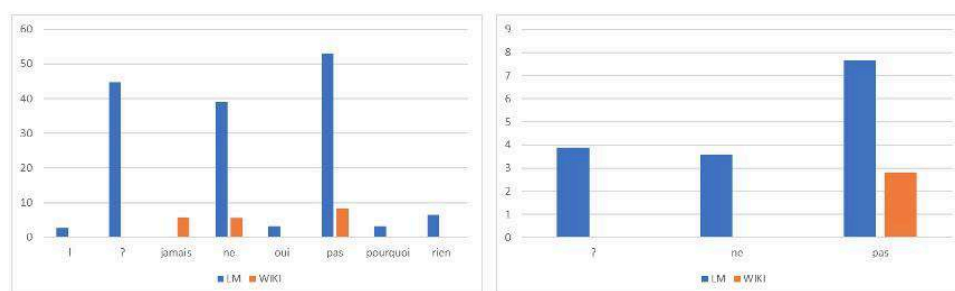
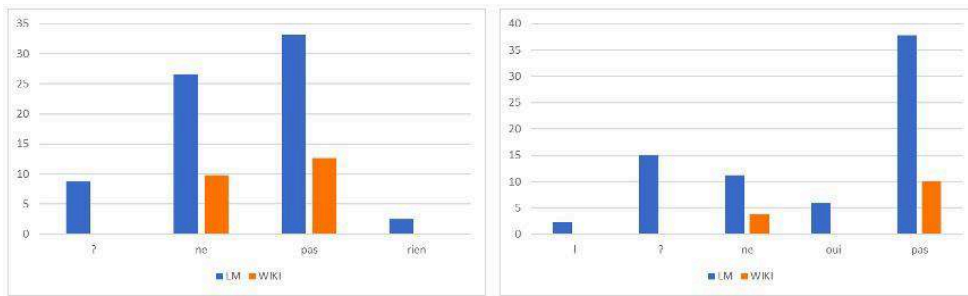


Figure 4 (à gauche) : Polyphonie type dialogique (*certes*)Figure 5 (à droite) : Polyphonie type dialogique (*certainement*)

- 46 Au niveau du nombre d'indices et de leur score, c'est *peut-être* qui manifeste la plus forte attraction pour ce type de marqueurs dans les deux corpus<sup>11</sup>. Il est caractérisé par des indications d'exclamations et d'interrogations, des marqueurs de négation et le marqueur d'approbation *oui* dans le corpus de presse. *Probablement* est l'adverbe qui attire le moins ce type d'expressions ; *certainement* et *certes* sont à mi-chemin entre *peut-être* et *probablement* avec le marqueur *oui*, qui caractérise *certainement* et non *certes*. Il est possible que cela soit dû au fait que *certainement* sert également de marqueur d'approbation alors que *certes* tend à perdre cet emploi.
- 47 On trouve en effet ce type d'emploi uniquement pour *certainement* et *peut-être*, dans le corpus *Le Monde* :

[15]

— Les analystes ont plutôt insisté sur l'abstention de la clientèle traditionnelle de la droite...

C. Braconnier : **Oui**, il y a très **certainement** eu une abstention plus forte que d'habitude de la part d'électeurs déçus par l'action de Nicolas Sarkozy, notamment dans les quartiers populaires où le président de la République était parvenu à capter une fraction de l'électorat sur les questions de l'immigration et de la sécurité. (*Le Monde*, 19-03-2010)

[16] « Quand je l'ai raccompagné à l'aéroport, confie Pierre Pica, je lui ai une fois de plus posé la question. Il m'a répondu : "**oui**, il y a **peut-être** là un problème spécifique à la France. Mais je n'en sais rien". » (*Le Monde*, 04-06-2010)

- 48 À noter que, dans cet exemple, *peut-être* permet à la fois d'approuver et de concéder, avec le segment qui suit introduit par *mais*.

### 2.3.2. Résultats pour famille 2 (indices de la présence du locuteur)

Figure 6 (à gauche) : Présence du locuteur (*peut-être*)  
 Figure 7 (à droite) : Présence du locuteur (*probablement*)

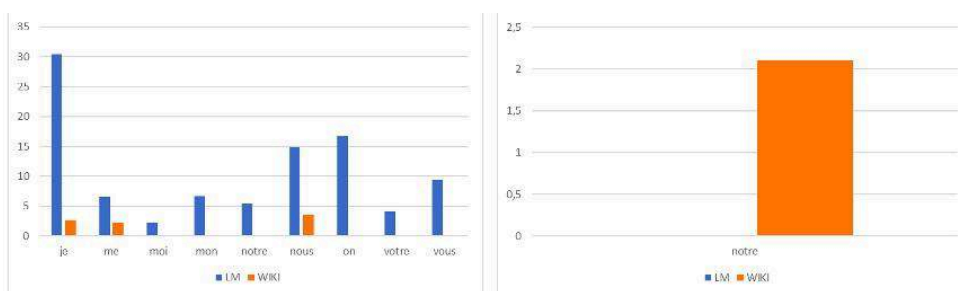


Figure 8 (à gauche) : Présence du locuteur (*certes*)  
 Figure 9 (à droite) : Présence du locuteur (*certainement*)



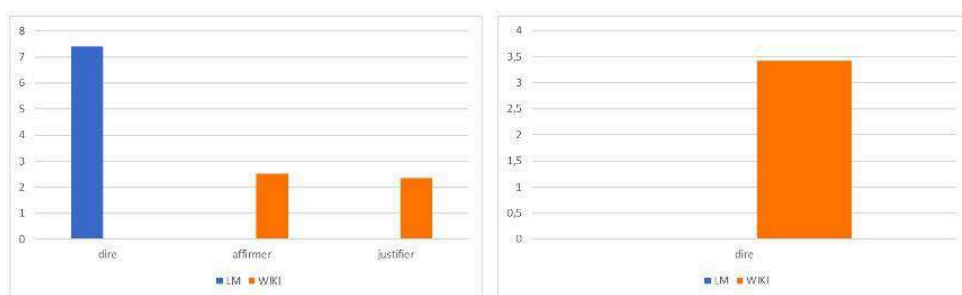
- 49 Dans cette catégorie, c'est à nouveau *peut-être* qui attire le plus de formes montrant que l'énoncé est susceptible d'être considéré comme *adressé* ; les autres adverbes en attirent nettement moins, avec un léger avantage pour *certainement* qui apparaît, comme *peut-être*, avec un pronom de 2<sup>e</sup> personne. Quand l'adverbe est utilisé pour modaliser le verbe à la 2<sup>e</sup> personne (*vous pensez/estimez... certainement/peut-être que...*), le caractère polyphonique de la tournure ressort explicitement puisqu'il s'agit d'attribuer un point de vue à un destinataire.

[17] De ses prunelles jaillissent alors des flammes, pas le moindre pompier catalan à l'horizon, Curzio Malaparte fixe l'ami, fusille l'amante, glacial : « Quoi ! Vous **croyez peut-être** que je ne vois pas votre manège ? Là ! Sous mes yeux ! Mon ami, mon amante ! Misérables ! (*Le Monde*, 24-03-2010)

[18] On me juge, en fait, pour un film dont moins d'un tiers était tourné au moment de mon arrestation. Vous **connaissez certainement** l'expression qui dit que ne dire que la moitié de la phrase « il n'y a point de Dieu que Dieu le grand » est synonyme de blasphème. (*Le Monde*, 22-12-2010)

### 2.3.3. Résultats pour la famille 3 (verbes de dire)

Figure 10 (à gauche) : Verbes de dire (*peut-être*)  
 Figure 11 (à droite) : Verbes de dire (*probablement*)



50 La catégorie ‘verbes de dire’ met en évidence *peut-être* uniquement avec trois verbes : *dire* dans *Le Monde*, et *affirmer* et *justifier* dans Wikipédia. *Certes* et *certainement* n’en présentent aucun qui leur serait spécifique et *probablement* est attiré par *dire* dans Wikipédia. Deux usages doivent être distingués : ceux où l’adverbe modalise directement le verbe *dicendi* [19] et ceux où l’adverbe intervient dans une citation introduite par un verbe *dicendi*, il ne porte donc pas sur ce dernier [20]. Ce dernier emploi est largement majoritaire, mais on trouve quelques occurrences du premier dans *Le Monde* avec *peut-être* uniquement.

[19] Paradoxalement, l’Histoire **dira peut-être** que ce fut une bonne journée pour l’Europe. (*Le Monde*, 13-02-2010)

[20] Robert Miklitsch **affirme** que la « téléphilie » de Tarantino est **peut-être** plus importante dans la sensibilité guidant *Pulp Fiction* que l’amour du réalisateur pour le rock’n’roll [...]. (Wikipédia 2019)

### 2.3.4. Résultat pour la famille 4 (opérateurs argumentatifs)

Figure 12 (à gauche) : Opérateurs argumentatifs (*peut-être*)  
 Figure 13 (à droite) : Opérateurs argumentatifs (*probablement*)

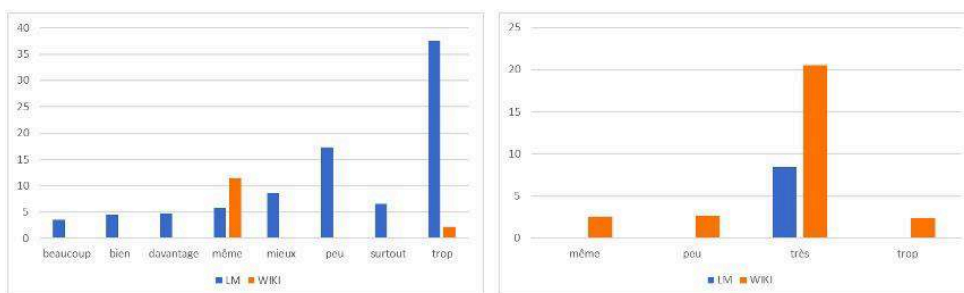
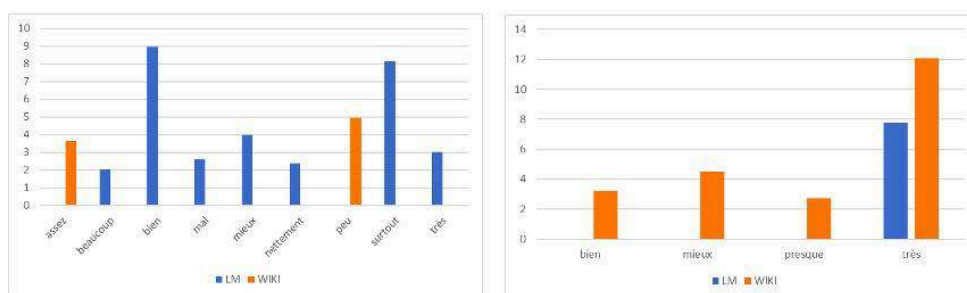


Figure 14 (à gauche) : Opérateurs argumentatifs (*certes*)  
 Figure 15 (à droite) : Opérateurs argumentatifs (*certainement*)



- 51 Les opérateurs argumentatifs montrent une dichotomie entre *peut-être* et *certes* d'un côté et *certainement* et *probablement* de l'autre. Les deux premiers adverbes comptent une variété nettement plus importante de marqueurs que les deux derniers (8 pour *peut-être* et 9 pour *certes* contre 4 pour *probablement* et *certainement*). Cette différence concerne également le type d'emploi de ces marqueurs : ils concernent l'énoncé dans lequel *peut-être* et *certes* sont utilisés, alors qu'ils peuvent concerner l'adverbe même en servant de spécificateur pour *probablement* et *certainement*. On trouve des emplois du type : *très probablement/presque certainement* :

[21] Il existe une tradition — **très probablement** erronée — selon laquelle les manuscrits contenus dans la bibliothèque de la ville auraient brûlé à ce moment. (Wikipédia 2019)

[22] [...] les efforts de la Chine vont **presque certainement** contribuer à améliorer les systèmes spatiaux militaires. (Wikipédia 2019)

- 52 alors qu'avec *peut-être*, l'opérateur *peu* porte sur l'orientation argumentative même de l'énoncé :

[23] M. Sarkozy n'a pas le choix, tant il a misé sur la réforme des retraites, mère des batailles qui doit permettre de réévaluer à la hausse son bilan et donner de lui l'image d'un homme à la personnalité **peut-être peu** aimée mais qui aura réformé jusqu'à la fin du quinquennat. (*Le Monde*, 21-10-2010)

### 2.3.5. Synthèse

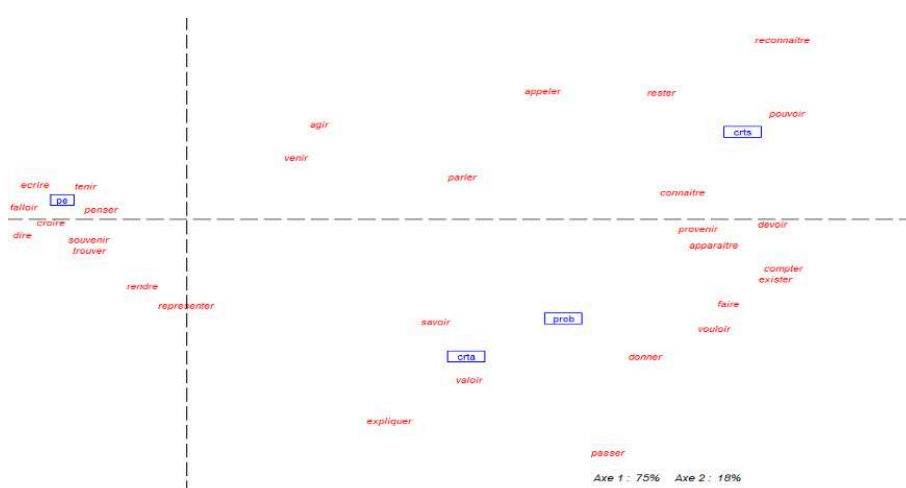
- 53 Les graphiques fondés sur les indices de spécificité concernant les quatre familles de cooccurrents montrent une différence nette entre *peut-être* et les trois autres adverbes : *peut-être* est caractérisé par son attrait plus marqué avec ces quatre familles de formes. Cette propriété est cohérente avec son PU, qui est lié à son potentiel dialogique. *Certes* se singularise quant à lui surtout par rapport aux marqueurs d'orientation argumentative (cas où il rejoint une utilisation semblable à *peut-être*). Cela est également cohérent avec sa propriété d'accord avec un contenu qui lui donne un PU propice aux environnements concessifs (donc à forte dimension argumentative). Enfin, il faut relever que *certainement* paraît avoir un comportement qui, pour certains marqueurs, se situe davantage du côté de *certes* ou *peut-être* et, pour d'autres, du côté de *probablement* : on a vu qu'il partage plus de marqueurs avec *certes* et *peut-être* pour les marqueurs de type polyphonie dialogique, qu'il partage le pronom de deuxième personne avec *peut-être*, mais qu'il est plus proche de *probablement* pour les opérateurs argumentatifs.

- 54 Une étude statistique fondée sur les co-fréquences permet de voir au moyen d'une autre méthode comment *certainement* se situe par rapport à ces trois adverbes.

## 2.4. Étude des co-fréquences représentée par une analyse factorielle des correspondances

- 55 La méthode que nous allons utiliser dans cette partie permet de saisir les différences entre chaque adverbe sur la base de leur comportement avec une série de formes données. Nous avons constitué un corpus regroupant les occurrences de ces quatre adverbes dans un empan gauche et droit de 5 items dans les quatre corpus (Wikipédia, *L'Est Républicain*, *Le Monde* et les discours des Présidents)<sup>12</sup>. Cela représente un total d'environ 40 000 *tokens*. À partir de ce corpus, nous avons extrait les 30 verbes les plus fréquents<sup>13</sup> (sans les auxiliaires *être*, *avoir* et le verbe *aller*) et calculé leur co-fréquence avec chacun des adverbes. Les résultats sont représentés sous forme d'une Analyse Factorielle des Correspondances (AFC) réalisée avec le logiciel Hyperbase (Brunet, 2012), qui reproduit au moyen d'un tableau de données la distance entre les quatre adverbes en fonction de leurs co-fréquences avec ces verbes. Chaque axe de l'AFC synthétise un pourcentage de cette information.

Figure 16 : AFC des trente verbes les plus fréquents se trouvant en situation de co-fréquence avec les quatre adverbes



- 56 Les tendances qui particularisent *peut-être* par rapport aux autres adverbes sont confirmées et accentuées au moyen de cette représentation. Ce dernier est tiré sur l'axe 1 (horizontal), axe qui reproduit 75 % de l'information, par des verbes renvoyant à l'acte d'énonciation (*dire, écrire*) ainsi que par des verbes renvoyant à une activité de pensée (*croire, penser, souvenir*). *Certes* se trouve sur l'autre côté de cet axe, attiré notamment par un verbe qui renvoie au processus propre à la concession (*reconnaître*). *Certainement* et *probablement* sont rapprochés et se situent à proximité de verbes liés à la connaissance (*expliquer, savoir*). On voit ainsi se dessiner une proximité entre *probablement* et *certainement*, difficilement identifiable avec les indices de cooccurrences, alors qu'elle ressortait avec les tests concernant le choix des locuteur-trices (cf. 2.2).

### 3. Synthèse et conclusions

- 57 À partir d'une analyse qualitative fondée sur les propriétés énonciatives des quatre adverbes, on a établi un profil sémantique qui permet de différencier *peut-être* par une attitude d'« accord » caractérisant tant le dire que le dit, *certaines* par une attitude d'« accord » caractérisant le dit, et *certainement* et *probablement* par une attitude de « prise en charge » caractérisant le dit. Ce profil a été mis en lien avec un PU que nous avons testé au moyen de deux méthodes : une méthode fondée sur les compétences linguistiques de francophones qui a montré que le choix d'un de ces adverbes dans un contexte donné n'est pas aléatoire, et différentes méthodes statistiques qui ont montré que le PU envisagé est cohérent avec la façon dont sont utilisés ces adverbes dans les corpus pris en compte. On a relevé que les adverbes caractérisés par une indication d'« accord » sont privilégiés dans les corpus qui ont une composante argumentative assumée ou revendiquée. Les indices de spécificité ont particularisé *peut-être*, qui est le seul à attirer les quatre familles d'indications montrant une polyphonie ou une composante argumentative. Enfin, la méthode fondée sur le calcul des cooccurrences a permis de mieux situer ces adverbes les uns par rapport aux autres : les regroupements font apparaître *certainement* et *probablement* comme partageant plus de formes que *peut-être* et *certaines*, qui se sont situés sur des pôles opposés de l'axe 1.
- 58 Au-delà de la possibilité de mieux circonscrire les caractéristiques sémantiques et d'usage de chacun de ces adverbes, notre démarche vise à la mise en place d'une méthodologie fondée sur la complémentarité des approches qualitatives avec les approches empiriques (utilisant des tests d'usage ou des données statistiques sur l'usage effectif des marqueurs dans les corpus). Notre étude donne des bases modulables en fonction de la nature de l'objet de recherche avec des possibilités d'élargissement en termes de genres discursifs et de langues.

---

### BIBLIOGRAPHIE

- ABEILLÉ A. & GODARD D. (dir.), 2021, *Grande Grammaire du français*, Arles, Actes Sud.
- ADAM J.-M., 1997, « Du renforcement de l'assertion à la concession : variations d'emploi de *certaines* », *L'Information grammaticale* 73, 3-9.
- ANSCOMBRE J.-C., 1981, « Marqueurs et hypermarqueurs de dérivation illocutoire : notions et problèmes », *Cahiers de linguistique française* 3, 75-124.
- BIBER D., JOHANSSON S., LEECH G., CONRAD S. & FINEGAN E., 1999, *The Longman Grammar of Spoken and Written English*, London, Longman.
- BORILLO A., 1976, « Les adverbes et la modalisation de l'assertion », *Langue française* 30, 74-89.
- BRUNET E., 2012, « Nouveau traitement des cooccurrences dans Hyperbase », *Corpus* 11, 219-246.  
DOI : 10.4000/corpus.2275

- CAREL M., 2010, « Polyphonie et évidentialité », manuscrit, *Semanticar* [en ligne], 1-33.  
URL : <https://semanticar.hypotheses.org/files/2024/09/carel-evidentialite.pdf>
- DE CESARE A.-M., 2022, « 20 Adverbes modaux et évidentiels », in G. Haßler (dir.), *Manuel des modes et modalités*, Berlin, Boston, De Gruyter, 507-532.
- DUCROT O., 2001, « Quelques raisons de distinguer “locuteur” et “énonciateurs” », *Polyphonie – Linguistique et littéraire* 3, 1-17.
- DUCROT O., 1989, *Logique, structure, énonciation. Lectures sur le langage*, Paris, Minuit.
- DUCROT O., 1984, *Le Dire et le dit*, Paris, Minuit.
- DUCROT O., 1983, « Opérateurs argumentatifs et visée argumentative », *Cahiers de Linguistique française* 5, 7-36.
- DUCROT O., 1980, *Les échelles argumentatives*, Paris, Minuit.
- GIL A., 1995, *Textadverbialen in den romanischen Sprachen: eine integrale Studie zu Konnektoren und Modalisatoren im Spanischen, Französischen und Italienischen*, Frankfurt am Main et al., Lang.
- GOSSELIN L., 2010, *Les modalités en français : la validation des représentations*, Amsterdam, New York, Rodopi.
- HEIDEN S., MAGUÉ J.-P. & PINCEMIN B., 2010, « TXM : Une plateforme logicielle open-source pour la textométrie – conception et développement », *10<sup>th</sup> International Conference on the Statistical Analysis of Textual Data – JADT 2010*, Jun. 2010, Rome, Italie, 1021-1032.
- HUDDLESTON R. & PULLUM G. K., 2002, *The Cambridge Grammar of the English Language*, Cambridge, Cambridge University Press.
- LAFON P., 1980, « Sur la variabilité de la fréquence des formes dans un corpus », *Mots* 1, 127-165.
- LEBLANC J.-M., FLEURY S. & NÉE E., 2017, « Quels outils logiciels et pour quoi faire ? », in E. Née (dir.) *Méthodes et outils informatiques pour l'analyse des discours*, Rennes, Presses universitaires de Rennes, 123-161.
- MOLINIER C. & LEVRIER F., 2000, *Grammaire des adverbes. Description des formes en -ment*, Genève, Droz.
- NUYTS J., 2014, “Analyses of the modal meanings”, in J. Nuyts & J. van der Auwera (dir.), *The Oxford Handbook of modality and mood* [en ligne], 31-49. DOI : <https://doi.org/10.1093/oxfordhb/9780199591435.013.1>
- PAMUKSAÇ A. & SANVIDO L., 2023, « Distinguer des genres textuels en français langue étrangère : étude textométrique de marqueurs subjectifs », *Synergies Espagne* 16, 129-145.
- ROSSARI C., MONTRICHARD C., RICCI C. & SANVIDO L. (à paraître), « Le sens post-modal de *pouvoir* et *peut-être* : une approche textométrique », *Canadian Journal of Linguistics / Revue canadienne de linguistique : More than modal? Intersubjective and discursive dimensions of (post)modal expressions / Plus que modal ? Dimensions intersubjectives et discursives des expressions (post)modales*.
- QUIRK R., GREENBAUM S., LEECH G. & SVARTVIK J., 1985, *A Comprehensive Grammar of the English Language*, London, Longman.
- SIMON-VANDENBERGEN A.-M. & AIJMER K., 2007, *The Semantic Field of Modal Certainty: A Corpus-Based Study of English Adverbs*, Berlin, New York, De Gruyter Mouton.
- TOGNINI-BONELLI E., 2001, *Corpus Linguistics at Work*, Amsterdam, J. Benjamins.

VOLD E. T., 2008, *Modalité épistémique et discours scientifique. Une étude contrastive des modalisateurs épistémiques dans des articles de recherche français, norvégiens et anglais, en linguistique et médecine*, Thèse de doctorat en philosophie, Université de Bergen.

WILMET M., 2003, *Grammaire critique du français*, 3<sup>e</sup> éd, Louvain-la-Neuve, Duculot.

## NOTES

1. Cet article est une contribution au projet *L'ancrage argumentatif des formes modales. Étude sur corpus avec un éclairage comparatif entre français et italien* financé par le FNS (projet n° 10001F\_192247), ainsi qu'au projet *Mobilettrés\_2024* qui vise la préparation de corpus recouvrant différents genres discursifs. Nous remercions les évaluateurs-trices pour leur relecture attentive et leurs commentaires très pertinents sur une première version de ce travail.

2. Nous utilisons un langage inclusif. Quand le terme « locuteur » renvoie à un concept théorique, nous utilisons la dénomination adoptée dans le cadre dont il est issu.

3. Les corpus interrogés sont Wikipédia 2019 (18 715 455 *tokens*) pour le genre encyclopédique, *Le Monde* 2010 (17 895 009 *tokens*) et *l'Est Républicain* 2010 (27 227 927 *tokens*) pour le genre presse. Pour le genre politique, nous avons utilisé le corpus « Discours présidentiels français de 1958 à aujourd'hui », <http://hyperbase.unice.fr/hyperbase/> (2 843 582 *tokens*). Nous disposons également d'un corpus partitionné selon le genre regroupant les discours présidentiels français, un échantillon du corpus Wikipédia (3 754 205 *tokens*) et un échantillon du corpus *Le Monde* (2'029'701 *tokens*). Tous nos corpus ont été lemmatisés avec TreeTagger et sont exploités à partir de la plateforme TXM (Heiden *et al.*, 2010).

4. Notre traduction.

5. Suivant la distinction opérée par Tognini-Bonelli (2001), nous nous situons dans une démarche mêlant à la fois une perspective *corpus-based*, dans la mesure où nous partons d'une hypothèse théorique pour conduire nos recherches sur corpus, et une perspective *corpus-driven*, dans la mesure où, d'une part, nous n'utilisons pas le corpus comme un réservoir d'exemples, mais nous l'utilisons pour mesurer les contextes privilégiés des formes que nous analysons, et d'autre part, à partir des résultats obtenus sur le corpus, nous revenons sur l'hypothèse théorique de départ qui définit le profil sémantique de la forme en cause.

6.  $T = [(3-4,9)^2 + (22 - 4,9)^2 + \dots + (2 - 4,9)^2] / 4,9$ .  $T > 23,59$  implique un rejet de l'indépendance avec une confiance supérieure à 99,5 %. Phrase 1 :  $T = 106,7$  ; phrase 2 :  $T = 84,08$  ; phrase 3 :  $T = 114,87$ .

7. Pour la phrase 2, un des sujets n'a pas répondu (d'où le total de 48 au lieu de 49), mais cela n'impacte pas le rejet à l'indépendance.

8. Nous renvoyons à Lafon (1980) pour une indication concernant le calcul probabiliste derrière cet indice.

9. Un test effectué en faisant varier les empan mène au constat qu'un empan deux fois plus petit ne permet pas de récupérer des cooccurrences pertinents tels que des adverbes dont la portée peut être relativement large. Toutefois, avec un empan deux fois plus large, on observe des résultats en contexte révélant un trop grand nombre d'occurrences des formes où celles-ci ne sont plus dans la portée de l'adverbe.

10. Pour la notion de *contenu adressé*, nous reprenons l'explication de Carel (2010 : 26) : « Je dirai qu'un contenu est *fortement adressé* lorsque le locuteur s'adresse à un interlocuteur actif, c'est-à-dire lorsque le locuteur attend ou du moins accepte une réponse. » Dans ce cadre, nous postulons qu'un énoncé en *je* mettrait plus en avant le caractère *adressé* de son contenu qu'un énoncé en *il*.

11. Nous présentons les résultats relatifs à deux de nos corpus pour des questions de place. Nous avons sélectionné les corpus qui font le plus apparaître de données et qui maintiennent une

différence en termes de genres, ce qui permet d'évaluer le caractère plus ou moins stable des cooccurrents que l'adverbe attire.

12. L'AFC permet de voir la proximité entre les quatre adverbes sans chercher à les différencier en fonction des genres, raison pour laquelle nous avons utilisé les quatre corpus pour avoir le plus de données possible.

13. Comme il s'agit des verbes, nous avons sélectionné un empan plus étroit que celui sélectionné dans la partie 2.3.

---

## RÉSUMÉS

Les adverbes signalant une posture épistémique sont considérés comme faisant partie de la catégorie des adverbes modaux. Ils se caractérisent par leur propriété de qualifier la valeur de vérité d'une proposition. Notre propos sera d'établir une méthodologie fondée sur une approche associant étroitement une démarche théorique mobilisant un cadre de linguistique énonciative et une démarche empirique recourant à des tests d'utilisation et à l'extraction des sites d'apparition de ces adverbes dans des corpus représentant différents genres discursifs. Nous verrons que ces deux méthodes donnent des résultats convergents qui confirment les caractéristiques issues de leur analyse énonciative.

Adverbs indicating an epistemic stance are considered to be part of the category of modal adverbs. They are usually characterized by their ability to qualify the truth value of a proposition. Our aim will be to establish a methodology closely associating the theoretical framework of enunciative linguistics and an empirical approach employing usage tests by native speakers and statistical tools to extract the specific linguistic environments of these adverbs in corpora representing different discursive genres. We will see that these two methods yield convergent results confirming the characteristics derived from their enunciative analysis.

## INDEX

**Mots-clés :** modalité, adverbes épistémiques, linguistique énonciative, linguistique de corpus

**Keywords :** modality, epistemic adverbs, enunciative linguistics, corpus linguistics

## AUTEURS

**CORINNE ROSSARI**

Université de Neuchâtel

**LAURA AUBRY**

Université de Neuchâtel